International Encyclopedia of Political Science

Bertrand Badie

International Encyclopedia of

POLITICAL SCIENCE

VOLUME 1

Editorial Board

International Encyclopedia of POLITICAL SCIENCE

VOLUME 1

EDITORS

BERTRAND BADIE
Sciences Po, Paris, France

DIRK BERG-SCHLOSSER
Philipps-Universität Marburg, Germany

LEONARDO MORLINO
LUISS Guido Carli, Italy

SAGE reference

Los Angeles | London | New Delhi
Singapore | Washington DC

IPSA AISP

Developed in partnership with the
International Political Science Association

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
33 Pekin Street #02-01
Far East Square
Singapore 048763

Publisher: Rolf A. Janke
Assistant to the Publisher: Michele Thompson
Developmental Editor: Diana E. Axelsen
Production Editor: Tracy Buyan
Reference Systems Manager: Leticia Gutierrez
Reference Systems Coordinator: Laura Notton
Copy Editor: QuADS Prepress (P) Ltd.
Typesetter: C&M Digitals (P) Ltd.
Proofreaders: Kate Peterson, Olivia Weber-Stenis
Indexer: Mary Mortensen
Cover Designer: Janet Kiesel
Marketing Manager: Kristi Ward

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

International encyclopedia of political science / edited by Bertrand Badie, Dirk Berg-Schlosser, Leonardo Morlino.

8 v. cm.
A SAGE Reference publication.
Includes bibliographical references and index.

ISBN 978-1-4129-5963-6 (cloth)

1. Political science—Encyclopedias. I. Badie, Bertrand.
II. Berg-Schlosser, Dirk. III. Morlino, Leonardo

JA61.I58 2011
320.03—dc22   2011009801

11 12 13 14 15 10 9 8 7 6 5 4 3 2 1

Contents

1

List of Entries

Reader's Guide

Comparative Politics, Theory, and Methods

Anarchism
Anarchy
Breakdown of Political Systems
Cabinets
Censorship
Central Banks
Change, Institutional
Charisma
Citizenship
Collaboration
Comparative Methods
Comparative Politics
Competition, Political
Conditionality
Constitutional Engineering
Corporativism
Decentralization
Democracy, Types of
Discursive Institutionalism
Elites
Environmental Issues
Executive
Government
Historical Sociology
Human Rights, Comparative Perspectives
Hybrid Regimes
Institutionalization
Institutions and Institutionalism
Interest Groups
Irredentism
Labor Movement
Leadership
Legitimacy
Military Rule
Monarchy
Neo-Patrimonialism
Neo-Weberian State
Oligarchy
Path Dependence
Personalization of Politics
Pillarization
Political Integration
Political Science, International Institutionalization
Political Systems, Types
Politics of Language
Presidentialism
Prospect Theory
Qualitative Comparative Analysis
Referenda
Reform
Regime (Comparative Politics)
Regionalism
Regionalization
Representation
Republic
Republicanism
Responsibility
Responsiveness
Revolution
Rule of Law
Secession
Semipresidentialism
Separation of Powers
Social Movements
Socialist Systems
Stability
State
State, Virtual
Terrorist Groups
Totalitarian Regimes
Welfare Policies
Welfare State

Case and Area Studies

Area Studies
Authoritarian Regimes

Subnational Governments

Democracy and Democratization

Decision Making in Democracies

Epistemological Foundations

Gender and Race/Ethnicity

International Relations

Peace, War, and Conflict Resolution

Political Economy

Political Parties

Political Philosophy/Theory

Equality and Inequality

Formal and Positive Theory

Theorists

Political Sociology

Public Policy and Public Administration

Public Policy

Policy Formulation
Policy Framing
Policy Instruments
Policy Learning
Policy Network
Policy Process, Models of
Prospect Theory
Reorganization
Risk and Public Policy
Self-Regulation
Soft Law
Stages Model of Policy Making
Think Tanks
Tragedy of the Commons
Transaction Costs

Public Administration

Administration
Administration Theory
Audit Society
Auditing
Autonomy, Administrative
Budgeting, Rational Models
Bureaucracy
Bureaucracy, Rational Choice Models
Bureaucracy, Street-Level
Civil Service
Corruption, Administrative
Effectiveness, Bureaucratic
Governance
Governance, Administration Policies
Governance, Informal
Governance, Multilevel
Governance, Urban
Governance Networks
Groupthink
Health Policy
Intelligence
Pay for Performance
Performance
Performance Management
Planning
Police
Politicization of Bureaucracy
Politicization of Civil Service
Public Budgeting
Public Employment
Public Goods
Public Office, Rewards
Regulation
Representative Bureaucracy
Responsiveness of Bureaucracy
Secret Services
Security Apparatus

Qualitative Methods

Analytic Narratives: Applications
Analytic Narratives: The Method
Configurational Comparative Methods
Data, Textual
Discourse Analysis
Ethnographic Methods
Evaluation Research
Fuzzy-Set Analysis
Grounded Theory
Hermeneutics
Interviewing
Interviews, Elite
Interviews, Expert
Mixed Methods
Network Analysis
Participant Observation
Process Tracing
Qualitative Comparative Analysis
Quantitative Versus Qualitative Methods
Thick Description
Triangulation

Quantitative Methods

Aggregate Data Analysis
Analysis of Variance
Boolean Algebra
Categorical Response Data
Censored and Truncated Data
Cohort Analysis
Correlation
Correspondence Analysis
Cross-National Surveys
Cross-Tabular Analysis
Data, Archival
Data, Missing
Data, Spatial
Data Analysis, Exploratory
Data Visualization
Event Counts
Event History Analysis
Experiments, Field

Religion

About the General Editors

Bertrand Badie is Director of Graduate Studies in International Relations and Professor of Political Science and International Relations in Sciences Po, Paris. He is also a research fellow at the Center for International Studies and Research (Centre d'études et de recherches internationales [CERI], Paris). He is currently working on international systems, diplomacy, multilateralism, and the sociology of international relations, after a long period devoted to research on the sociology of the state. He has been a visiting professor in many European, Asian, and African universities. He started his academic career at Université of Paris 1 Panthéon-Sorbonne (1974–1982), followed by a stint at Clermont-Ferrand University (1982–1990). He has served as Director of Sciences Po University Press (1994–2003), a member of the International Political Science Association (IPSA) Executive Committee (2003–2009), and Vice President of IPSA (2006–2009). He has published about 20 books—translated into 15 languages—including *The Sociology of the State* (1983), with Pierre Birnbaum; *The Imported State* (2000); *Les Deux Etats* (The Two States; 1987); *La Fin des Territoires* (The End of Territories; 1995); *Un Monde sans souveraineté* (A World Without Sovereignty; 1999); *La Diplomatie des Droits de l'Homme* (The Diplomacy of Human Rights; 2002); *L'Impuissance de la Puissance* (The Impotence of Power; 2004); *Le Diplomate et l'Intrus* (The Diplomat and the Intruders; 2008); and *La Diplomatie de Connivence* (The Diplomacy of Collusion; 2011). He earned his Ph.D. in political science at Sciences Po (1975) and also holds master's degrees in law (1972), in contemporary history (1977), and in oriental studies (1975). He is "Agrégé en science politique" (1982). He also holds an honorary Ph.D. from FUCaM (Facultés Universitaires Catholiques de Mons), in Mons, Belgium.

Dirk Berg-Schlosser is Professor Emeritus at Philipps-University, Marburg, Germany, where he has also been Director of the Institute of Political Science and Dean of the Faculty of Social Sciences and Philosophy. He has conducted research and taught at the universities of Munich, Aachen, Augsburg, and Eichstaett (Germany); Nairobi (Kenya); Stellenbosch (South Africa); Sciences Po (Paris); and California at Berkeley. He was chairman of the Research Committee on "Democratization in Comparative Perspective" of the International Political Science Association (IPSA) (1988–2000). He has served as Chair of the European Consortium for Political Research (ECPR) (2003–2006) and as Vice-President of IPSA (2006–2009). He is currently director of the IPSA Summer Schools on Research Methods at the University of São Paulo and Stellenbosch University. His major research interests include empirical democratic theory, development studies, political culture, comparative politics, and comparative methodology. He is author and coauthor of more than 20 books and about 200 articles in major international journals and chapters in edited volumes. His major books in English include *Tradition and Change in Kenya* (1984); *Political Stability and Development* (1991), with Rainer Siegler; *Conditions of Democracy in Europe, 1919–39: Systematic Case Studies* (2000), with Jeremy Mitchell; *Authoritarianism and Democracy in Europe, 1919–39: Comparative Analyses* (2002), with Jeremy Mitchell; *Poverty and Democracy* (2003), with Norbert Kersting; and *Democratization: The State of the Art* (2007). He is a member of the board of six major international journals and coeditor of book series with Oxford University Press and Campus. He has been the organizer and chair of numerous national and international panels, workshops, symposia, and conferences. He has served as referee for numerous national and international funding organizations.

He is currently a member of the Prize Committee of the Johan-Skytte Prize and a fellow at the Stellenbosch Institute for Advanced Study (STIAS). He obtained his first doctorate (Dr. oec. publ.) in economics at the University of Munich, a second advanced degree (Dr. phil. habil.) in political science at the University of Augsburg, and a Ph.D. in political science from the University of California at Berkeley.

Leonardo Morlino is Professor of Political Science in the School of Government and Director of the Research Centre on Democracies and Democratizations at Libera Università Internazionale degli Studi Sociali (LUISS) Guido Carli, in Rome. In 2004, he was awarded the Jean Monnet Chair in European Studies by the European Commission. He is currently President of the International Political Science Association (IPSA) (2009–2012). He has been a visiting professor or fellow at Yale University, Stanford University, Nuffield College (Oxford, UK), Institut d'Etudes Politiques (Paris), Hoover Institution, Center for Advanced Research in Social Sciences (Madrid, Spain), and St. Anthony's College, Oxford. His main research interests include the methodology of comparison, the processes of democratization, and the role of parties, interest groups, and civil society in contemporary European democracies. Since 2003, he has been actively involved in research on the promotion of democracy and rule of law in the new neighbors of the European Union and on the qualities of democracy in Europe and Latin America. He has mainly contributed to theory on regime change and democratizations, with specific attention to the so-called theory of anchoring, the quality of democracy and its empirical analysis. He has published more than 30 books and about 200 chapters in books and articles in journals in Italian, French, English, Spanish, German, Hungarian, and Japanese—mainly on the methodology of comparison, authoritarianisms, the European Union, democratic theory, and democratization processes. His main books as author include *Changes for Democracy: Actors, Structures, Processes* (2011); *Democrazie e Democratizzazioni* (Democracy and Democratization, 2003); *Democracy Between Consolidation and Crisis: Parties, Groups, and Citizens in Southern Europe* (1998); and *Como Cambian los regimenes politicos* (How Political Regimes Change, 1985). Books that he has coedited and/or coauthored include *Democratization and the European Union: Comparing Central and Eastern European Post-Communist Countries* (2010); *International Actors, Democratization and the Rule of Law: Anchoring Democracy?* (2008); *Partiti e caso italiano* (Parties and the Italian Case; 2006); *Europeizzazione e rappresentanza territoriale: Il caso italiano* (Europeanization and Territorial Representation: The Italian Case; 2006), *Party Change in Southern Europe* (2007); and *Assessing the Quality of Democracy* (2005). Some of his most recent articles in journals include "Are There Hybrid Regimes? Or Are They Just an Optical Illusion?" in *European Political Science Review* (2009); "Europeanization and Territorial Representation in Italy" (coauthor) in *West European Politics* (2006); "Anchors, Anchoring and Democratic Change" in *Comparative Political Studies* (2005); "What Is a 'Good' Democracy?" in *Democratization* (2004); "'Good' and 'Bad' Democracies: How to Conduct Research Into the Quality of Democracy" in *Communist Studies and Transition Politics* (2004); "Consolidation Démocratique: La Théorie de l'Ancrage" (Democratic Consolidation: The Theory of Anchorage) in *Revue Internationale de Politique Comparée* (2001); and "Architectures constitutionelles et politiques democratiques en Europe de l'Est" (Architectures Constitutional and Democratic Policies in Eastern Europe) in *Revue Française de Science Politique* (2000). Some of his most recent chapters in books include "Political Parties" in *Democratization* (2009); "What Changes in South European Parties: A Comparative Introduction" in *Party Change in Southern Europe* (2007); "Introduction" to *Assessing the Quality of Democracy: Theory and Empirical Analysis* (2005); "Problematizing the Links Between Authoritarian Legacies and 'Good' Democracy" (coauthor) in *Democracy and Authoritarian Legacies in Southern Europe and Latin America* (2004); "The Three Phases of Italian Parties" in *Political Parties and Democracy* (2001); "Democracy, Southern European Style?" (coauthor) in *Parties, Politics, and Democracy in the New Southern Europe* (2001); and "Constitutional Design and Problems of Implementation in Southern and Eastern Europe" in *Democratic Consolidations in Eastern Europe* (2001).

Contributors

Joel D. Aberbach
University of California, Los Angeles
Los Angeles, California, United States

Attila Ágh
Corvinus University of Budapest
Budapest, Hungary

John Agnew
University of California, Los Angeles
Los Angeles, California, United States

Despina Alexiadou
University of Pittsburgh
Pittsburgh, Pennsylvania, United States

Lincoln Allison
University of Warwick
Coventry, United Kingdom

Nicholas Almendares
New York University
New York City, New York, United States

Karen J. Alter
Northwestern University
Evanston, Illinois, United States

David Altman
Pontificia Universidad Catolica de Chile
Santiago, Chile

Thomas Ambrosio
North Dakota State University
Fargo, North Dakota, United States

Dag Anckar
Åbo Akademi
Åbo, Finland

Zehra F. Kabasakal Arat
Purchase College, State University of New York
Port Chester, New York, United States

Luz Marina Arias
University of California, San Diego
La Jolla, California, United States

Leslie Elliott Armijo
Portland State University
Lake Oswego, Oregon, United States

Klaus Armingeon
University of Bern
Bern, Switzerland

Kai Arzheimer
University of Essex
Colchester, United Kingdom

Raul Atria
University of Chile
Santiago, Chile

Fulvio Attinà
University of Catania
Catania, Italy

Jun Ayukawa
Kwansei Gakuin University
Nishinomiya, Japan

Bertrand Badie
Institut d'études politiques (Sciences Po)
Paris, France

Delia Bailey
YouGov
Palo Alto, California, United States

David A. Baldwin
Princeton University
Princeton, New Jersey, United States

Jinny Bang
New York University
New York, New York, United States

Michael Banton
University of Bristol
Bristol, United Kingdom

Daniel Barbu
University of Bucharest
Bucharest, Romania

Jens Bartelson
Lund University
Lund, Sweden

Jérémie Barthas
University of Johannesburg
Johannesburg, South Africa

Viva Ona Bartkus
University of Notre Dame
Notre Dame, Indiana, United States

Stefano Bartolini
European University Institute
San Domenico di Fiesole, Italy

Robert H. Bates
Harvard University
Cambridge, Massachusetts, United States

Dario Battistella
Sciences Po Bordeaux
Pessac, France

Julien Bauer
Université du Québec à Montréal
Montréal, Québec, Canada

Andrea Baumeister
University of Stirling
Stirling, United Kingdom

Richard John Beardsworth
American University of Paris
Paris, France

Nathaniel Beck
New York University
New York City, New York, United States

David Beetham
University of Leeds
Leeds, United Kingdom

Daniel Béland
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

Richard Bellamy
University College London
London, United Kingdom

Ulises Beltrán
Centro de Investigación y Docencia Económicas
Mexico City, Mexico

Andrew Bennett
Georgetown University
Washington, D.C., United States

Li Bennich-Björkman
Uppsala University
Uppsala, Sweden

Kenneth Benoit
Trinity College
Dublin, Ireland

David Benson
University of East Anglia
Norwich, United Kingdom

Arthur Benz
Technische Universität Darmstadt
Darmstadt, Germany

Jacob Bercovitch
University of Canterbury
Christchurch, New Zealand

Dirk Berg-Schlosser
Philipps-Universität Marburg
Marburg, Germany

Manfred Max Bergman
University of Basel
Basel, Switzerland

Nancy Bermeo
Nuffield College, University of Oxford
Oxford, United Kingdom

Rajeev Bhargava
Centre for the Study of Developing Societies
Delhi, India

Chris J. Bickerton
University of Amsterdam
Amsterdam, Netherlands

Pierre Birnbaum
Université Paris 1
Paris, France

James Warner Björkman
Institute of Social Studies
The Hague, Netherlands

Jörg Blasius
University of Bonn, Institute for Political Science and Sociology
Bonn, Germany

Jens Blom-Hansen
Aarhus University
Aarhus, Denmark

Jean Blondel
European University Institute
San Domenico di Fiesole, Italy

Mark Blyth
Brown University
Providence, Rhode Island, United States

Frederick J. Boehmke
University of Iowa
Iowa City, Iowa, United States

Matthijs Bogaards
Jacobs University Bremen
Bremen, Germany

Arjen Boin
Louisiana State University
Baton Rouge, Louisiana, United States

Maitseo Bolaane
University of Botswana
Gaborone, Botswana

Chris W. Bonneau
University of Pittsburgh
Pittsburgh, Pennsylvania, United States

Tanja A. Börzel
Freie Universität Berlin
Berlin, Germany

Mathias Bös
Philipps-Universität Marburg
Marburg, Germany

Chiara Bottici
The New School for Social Research
New York City, New York, United States

Mark Bovens
Utrecht University School of Governance
Utrecht, Netherlands

Bidhya Bowornwathana
Chulalongkorn University
Bangkok, Thailand

Janet M. Box-Steffensmeier
Ohio State University
Columbus, Ohio, United States

András Bozóki
Central European University
Budapest, Hungary

Steven J. Brams
New York University
New York City, New York, United States

Patrick T. Brandt
University of Texas at Dallas
Richardson, Texas, United States

Reinhard Brandt
Philipps-Universität Marburg
Marburg, Germany

Michael Bratton
Michigan State University
East Lansing, Michigan, United States

Philippe Braud
Sciences Po
Paris, France

Karl Braun
Philipps-Universität Marburg
Marburg, Germany

David R. Brillinger
University of California, Berkeley
Berkeley, California, United States

Malena Britz
Swedish National Defence College
Stockholm, Sweden

Stephen G. Brooks
Dartmouth College
Hanover, New Hampshire, United States

Alison Brysk
University of California, Irvine
Irvine, California, United States

Hubertus Buchstein
University of Greifswald
Greifswald, Germany

Ian Budge
University of Essex
Colchester, United Kingdom

Xymena Bukowska
Institute of Political Studies, University of Warsaw
Warsaw, Poland

François Burgat
IFPO CNRS (Damascus)
Aix-en-Provence, France

Craig Calhoun
New York University
New York City, New York, United States

Mauro Calise
University of Naples Federico II
Naples, Italy

John Thomas Callaghan
University of Salford
Greater Manchester, United Kingdom

Randall Calvert
Washington University in St. Louis
St. Louis, Missouri, United States

Jean-Yves Calvez (Deceased)
Jesuit Faculties of Philosophy and Theology
Paris, France

James A. Caporaso
University of Washington
Seattle, Washington, United States

Edward G. Carmines
Indiana University, Bloomington
Bloomington, Indiana, United States

Roberto Cartocci
University of Bologna
Bologna, Italy

Brunella Casalini
University of Florence
Florence, Italy

Bruno Cautrès
CEVIPOF/Sciences Po
Paris, France

Frédéric Charillon
Université d'Auvergne–Clermont Ferrand 1
Paris, France

Simon Chesterman
New York University School of Law/National University of Singapore
Singapore

Anthony B. L. Cheung
Hong Kong Institute of Education
Tai Po, New Territories, Hong Kong

Peter T. Y. Cheung
University of Hong Kong
Pokfulam, Hong Kong

Wendy K. Tam Cho
University of Illinois at Urbana-Champaign
Champaign, Illinois, United States

Geeta Chowdhry
Northern Arizona University
Flagstaff, Arizona, United States

Tom Christensen
University of Oslo
Oslo, Norway

Dino P. Christenson
Boston University
Boston, Massachusetts, United States

Dag Harald Claes
University of Oslo
Oslo, Norway

William Roberts Clark
University of Michigan
Ann Arbor, Michigan, United States

Harold D. Clarke
University of Texas at Dallas
Richardson, Texas, United States

Kevin A. Clarke
University of Rochester
Rochester, New York, United States

Sara Clavero
Queen's University Belfast
Belfast, United Kingdom

Jean-Marc Coicaud
United Nations University
New York City, New York, United States

Dominique Colas
Institut d'Etudes Politiques de Paris (Sciences Po)
Paris, France

H. K. Colebatch
University of New South Wales
Sydney, Australia

Josep M. Colomer
Institute for Economic Analysis, Higher Council for Scientific Research
Barcelona, Spain

Ariel Colonomos
CERI-Sciences Po/CNRS
Paris, France

Maurizio Cotta
University of Siena
Siena, Italy

Michael Cox
London School of Economics
London, United Kingdom

Bruce Cronin
City College of New York
New York City, New York, United States

Lasse Cronqvist
Universität Trier
Trier, Germany

Jane Curry
Santa Clara University
Santa Clara, California, United States

Roland Czada
University of Osnabrueck
Osnabrueck, Germany

Carl Johan Dahlström
University of Göteborg
Göteborg, Sweden

Jean-Pascal Daloz
St. Anthony's College, University of Oxford
Oxford, United Kingdom

David Darmofal
University of South Carolina
Columbia, South Carolina, United States

Charles-Philippe David
Université du Québec à Montréal
Montréal, Québec, Canada

James W. Davis
University of St. Gallen
St. Gallen, Switzerland

Gjalt de Graaf
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

Miguel De Luca
Universidad de Buenos Aires / CONICET
Buenos Aires, Argentina

Gisèle De Meur
Université Libre de Bruxelles
Brussels, Belgium

Liliana De Riz
Universidad de Buenos Aires
Buenos Aires, Argentina

Jaap de Wilde
University of Groningen
Groningen, Netherlands

Henk Dekker
Leiden University
Leiden, Netherlands

Donatella della Porta
European University Institute
San Domenico di Fiesole, Italy

Yves Déloye
Université Paris 1 Panthéon-Sorbonne
Paris, France

Bruno Dente
Politecnico di Milano
Milan, Italy

Guillaume Devin
Institut d'Etudes Politiques de Paris (Sciences Po)
Paris, France

Ilvo Diamanti
Università di Urbino "Carlo Bo"
Urbino, Italy

Alain Dieckhoff
CERI-Sciences Po/CNRS
Paris, France

Elias Dinas
European University Institute
San Domenico di Fiesole, Italy

Michel Dobry
Université Paris 1 Panthéon-Sorbonne
Paris, France

Mattei Dogan (Deceased)
National Center for Scientific Research (CNRS)
Paris, France

Lex Donaldson
University of New South Wales
Sydney, New South Wales, Australia

Jean-Yves Dormagen
University of Montpellier
Montpellier, France

Michael Doyle
Columbia University
New York City, New York, United States

Jaap Dronkers
European University Institute
San Domenico di Fiesole, Italy

Frédérick Guillaume Dufour
Université du Québec à Montréal
Montréal, Québec, Canada

Andreas Duit
Stockholm University
Stockholm, Sweden

Craig Malcolm Duncan
Ithaca College
Ithaca, New York, United States

Carl W. Dundas
Dundas and Associates Ltd.
Furzton, Milton Keynes, United Kingdom

Thad Dunning
Yale University
New Haven, Connecticut, United States

Francis Dupuis-Déri
Université du Québec à Montréal
Montréal, Québec, Canada

Andreas Dür
University of Salzburg
Salzburg, Austria

Burkard Eberlein
York University
Toronto, Ontario, Canada

S. N. Eisenstadt (Deceased)
The Van Leer Jerusalem Institute
Jerusalem, Israel

Rainer Eisfeld
University of Osnabrueck
Osnabrueck, Germany

Zachary Elkins
University of Texas at Austin
Austin, Texas, United States

Andrew Ellis
International Institute for Democracy and Electoral Assistance (IDEA)
Stockholm, Sweden

Kjell Engelbrekt
Swedish National Defense College
Stockholm, Sweden

Diane Ethier
University of Montréal
Montréal, Quebec, Canada

Amitai Etzioni
The Institute for Communitarian Policy Studies
Washington, D.C., United States

Sergio Fabbrini
Trento University
Trento, Italy

Jürgen W. Falter
Johannes Gutenberg University Mainz
Mainz, Germany

Harvey B. Feigenbaum
The George Washington University
Washington, D.C., United States

Yale H. Ferguson
Rutgers University–Newark
Newark, New Jersey, United States

Maurizio Ferrera
Università degli Studi di Milano
Milan, Italy

Mónica Ferrín
European University Institute
Florence, Italy

Louis Fisher
Library of Congress
Washington, D.C., United States

Matthew Flinders
University of Sheffield
Sheffield, United Kingdom

Tuomas Forsberg
University of Tampere
Tampere, Finland

Laurent Fourchard
Institut d'Etudes Politiques de Bordeaux
Pessac, France

Mark N. Franklin
European University Institute
San Domenico di Fiesole, Italy

Oscar W. Gabriel
Universität Stuttgart
Stuttgart, Germany

Sean Gailmard
University of California, Berkeley
Berkeley, California, United States

Yvonne Galligan
Queen's University Belfast
Belfast, United Kingdom

Steffen Ganghof
University of Potsdam
Potsdam, Germany

Philipp Genschel
Jacobs University Bremen
Bremen, Germany

Emilio Gentile
Sapienza–Università di Roma
Rome, Italy

Alan S. Gerber
Yale University
New Haven, Connecticut, United States

John Gerring
Boston University
Boston, Massachusetts, United States

Julie Gervais
Triangle Centre for Political Research, CNRS
Lyon, France

Jeff Gill
Washington University in St. Louis
St. Louis, Missouri, United States

Garrett Glasgow
University of California, Santa Barbara
Santa Barbara, California, United States

Kristian Skrede Gleditsch
University of Essex
Colchester, United Kingdom

Giampietro Gobo
Università degli Studi di Milano
Milan, Italy

Laurent Goetschel
University of Basel
Basel, Switzerland

Leslie Friedman Goldstein
University of Delaware
Newark, Delaware, United States

Wyn Grant
University of Warwick
Coventry, United Kingdom

Davide Grassi
Università degli studi di Torino
Turin, Italy

Matthew Gratias
University of Southern California
Los Angeles, California, United States

Simon Green
Aston University
Birmingham, United Kingdom

Michael Greenacre
Universitat Pompeu Fabra
Barcelona, Spain

Robert Gregory
Victoria University of Wellington
Wellington, New Zealand

Michael Th. Greven
Universität Hamburg
Hamburg, Germany

Bernard Grofman
University of California, Irvine
Irvine, California, United States

A. J. R. Groom
University of Kent
Canterbury, United Kingdom

Jürgen R. Grote
Charles University in Prague
Prague, Czech Republic

Siba N. Grovogui
Johns Hopkins University
Baltimore, Maryland, United States

Gérard Grunberg
CNRS/Sciences Po Paris
Paris, France

Carlo Guarnieri
University of Bologna
Bologna, Italy

John G. Gunnell
State University of New York
Albany, New York, United States
and University of California, Davis
Davis, California, United States

Jeroen Gunning
Aberystwyth University
Aberystwyth, United Kingdom

Dingping Guo
Fudan University
Shanghai, China

Asha Gupta
University of Delhi
Delhi, India

Stefano Guzzini
Danish Institute for International Studies & Uppsala University
Copenhagen, Denmark

Peter M. Haas
University of Massachusetts Amherst
Amherst, Massachusetts, United States

Christian W. Haerpfer
University of Aberdeen
Aberdeen, United Kingdom

Stephan Haggard
University of California, San Diego
La Jolla, California, United States

Maarten Hajer
Universiteit van Amsterdam
Amsterdam, Netherlands

Mark Hallerberg
Hertie School of Governance
Berlin, Germany

Amr Hamzawy
Carnegie Endowment for International Peace
Beirut, Lebanon

Dominik Hangartner
University of California, Berkeley
Berkeley, California, United States

Russell Hardin
New York University
New York City, New York, United States

Erin Hartman
University of California, Berkeley
Berkeley, California, United States

Andreas Hasenclever
Universität Tübingen
Tübingen, Germany

Pierre Hassner
Sciences Po
Paris, France

Jeffrey Haynes
London Metropolitan University
London, United Kingdom

Reuven Y. Hazan
Hebrew University of Jerusalem
Jerusalem, Israel

Jens Heibach
Philipps-Universität Marburg
Marburg, Germany

Eva G. Heidbreder
Hertie School of Governance
Berlin, Germany

Jorge Heine
Centre for International Governance Innovation
Waterloo, Ontario, Canada

Wilhelm Heitmeyer
Bielefeld University
Bielefeld, Germany

Gunther Hellmann
Goethe-University Frankfurt / Main
Frankfurt am Main, Germany

Guy Hermet
Sciences Po
Paris, France

Johannes Herrmann
Justus-Liebig-Universität Gießen
Gießen, Germany

F. Daniel Hidalgo
University of California, Berkeley
Berkeley, California, United States

John Higley
University of Texas at Austin
Austin, Texas, United States

Christopher Hill
University of Cambridge
Cambridge, United Kingdom

Jonas Hinnfors
University of Göteborg
Göteborg, Sweden

Katherine Hite
Vassar College
Poughkeepsie, New York, United States

Stephen Hobden
University of East London
London, United Kingdom

Brian Hocking
Loughborough University
Leicestershire, United Kingdom

Bert Hoffmann
German Institute of Global and Area Studies
Hamburg, Germany

Ursula Hoffmann-Lange
University of Bamberg
Bamberg, Germany

James F. Hollifield
SMU
Dallas, Texas, United States

James Hollyer
New York University
New York City, New York, United States

John A. Hoornbeek
Kent State University
Kent, Ohio, United States

Adrienne Hosek
University of California, Berkeley
Berkeley, California, United States

David Patrick Houghton
University of Central Florida
Orlando, Florida, United States

Michael Howlett
Simon Fraser University
Burnaby, British Columbia, Canada

Wei Hu
Shanghai Jiaotong University
Shanghai, China

Merlijn van Hulst
Universiteit van Tilburg
Tilburg, Netherlands

Carlos Huneeus
Universidad de Chile
Santiago, Chile

Christian Hunold
Drexel University
Philadelphia, Pennsylvania, United States

Peter L. Hupe
Erasmus University Rotterdam
Rotterdam, Netherlands

Andrew Hurrell
University of Oxford
Oxford, United Kingdom

Serge Hurtig
International Political Science Association
Paris, France

Adrian Hyde-Price
University of Bath
Bath, United Kingdom

Goran Hyden
University of Florida
Gainesville, Florida, United States

Piero Ignazi
University of Bologna
Bologna, Italy

Mikhail Ilyin
MGIMO University
Moscow, Russian Federation

Ellen M. Immergut
Humboldt University Berlin
Berlin, Germany

Takashi Inoguchi
University of Niigata Prefecture
Tokyo, Japan

Tatsuo Inoue
University of Tokyo
Kokubunji, Tokyo, Japan

Sebastian Jäckle
University of Heidelberg
Heidelberg, Germany

Simon Jackman
Stanford University
Stanford, California, United States

John E. Jackson
University of Michigan
Ann Arbor, Michigan, United States

William G. Jacoby
Michigan State University
East Lansing, Michigan, United States

Christophe Jaffrelot
CERI-Sciences Po/CNRS
Paris, France

Patrick James
University of Southern California
Los Angeles, California, United States

Laura Janik
University of Northern Iowa
Cedar Falls, Iowa, United States

Ben Jann
ETH Zurich
Zurich, Switzerland

Lucien Jaume
Sciences Po
Paris, France

Kanishka Jayasuriya
Murdoch University
Murdoch, Western Australia, Australia

Fabien Jobard
Centre de Recherche Sociologique sur le Droit et les Institutions Pénales, CNRS
Guyancourt, France

Loch K. Johnson
University of Georgia
Athens, Georgia, United States

Rebecca E. Johnson
Acronym Institute for Disarmament Diplomacy
London, United Kingdom

Michael Johnston
Colgate University
Hamilton, New York, United States

Andrew Jordan
University of East Anglia
Norwich, United Kingdom

Grant Jordan
University of Aberdeen
Aberdeen, United Kingdom

Knud Erik Jørgensen
University of Aarhus
Aarhus, Denmark

Dirk Jörke
University of Greifswald
Greifswald, Germany

Dirk Junge
University of Mannheim
Mannheim, Germany

Max Kaase
Jacobs University Bremen
Bremen, Germany

Lutz Kaelber
University of Vermont
Burlington, Vermont, United States

Dirk Kaesler
Philipps-Universität Marburg
Marburg, Germany

Ben Kamis
Universität Tübingen
Tübingen, Germany

Devesh Kapur
University of Pennsylvania
Philadelphia, Pennsylvania, United States

Lauri Karvonen
Åbo Akademi
Åbo, Finland

Riva Kastoryano
CERI-Sciences Po/CNRS
Paris, France

Michael Keating
European University Institute
San Domenico di Fiesole, Italy

Luke Keele
Ohio State University
Columbus, Ohio, United States

Hans Keman
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

Norbert Kersting
Stellenbosch University
Stellenbosch, South Africa

Mark Kesselman
Columbia University
New York City, New York, United States

Pekka Kettunen
University of Jyväskylä
Jyväskylä, Finland

Sun-Chul Kim
Barnard College
New York City, New York, United States

Leo Kissler
Philipps-Universität Marburg
Marburg, Germany

Markus Klein
Leibniz Universität Hannover
Hannover, Germany

Oddbjørn Knutsen
University of Oslo
Oslo, Norway

Torbjørn L. Knutsen
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway

Ken Kollman
University of Michigan
Ann Arbor, Michigan, United States

Serge-Christophe Kolm
École des Hautes Études en Sciences Sociales
Paris, France

Thomas König
University of Mannheim
Mannheim, Germany

Petr Kopecký
Leiden University
Leiden, Netherlands

Hennie Kotzé
Stellenbosch University
Matieland, South Africa

Mark Kramer
Harvard University
Cambridge, Massachusetts, United States

Matthew H. Kramer
Cambridge University
Cambridge, United Kingdom

Algis Krupavičius
Kaunas University of Technology
Kaunas, Lithuania

Milja Kurki
Aberystwyth University
Aberystwyth, United Kingdom

Guy Lachapelle
Concordia University
Montréal, Québec, Canada

Justine Lacroix
Université Libre de Bruxelles
Brussels, Belgium

Stella Ladi
Panteion University
Athens, Greece

Per Lægreid
University of Bergen
Bergen, Norway

David A. Lake
University of California, San Diego
La Jolla, California, United States

Wai Fung Lam
University of Hong Kong
Pokfulam, Hong Kong

Todd Landman
University of Essex
Colchester, United Kingdom

Jan-Erik Lane
University of Freiburg
Freiburg, Germany

Victor Lapuente
University of Göteborg
Göteborg, Sweden

Henry Laurens
Collège de France
Paris, France

Sébastien Laurent
University of Bordeaux
Bordeaux, France

Michael Laver
New York University
New York City, New York, United States

David S. Law
Washington University in St. Louis
St. Louis, Missouri, United States

Kay Lawson
San Francisco State University
San Francisco, California, United States

Marc Lazar
Sciences Po
Paris, France

Jacques Le Cacheux
University of Pau and Sciences Po Paris
Paris, France

Richard Ned Lebow
Dartmouth College
Hanover, New Hampshire, United States

Jean Leca
Institut d'Etudes Politiques de Paris (Sciences Po)
Paris, France

Donna Lee
University of Birmingham
Birmingham, United Kingdom

Beth L. Leech
Rutgers University
New Brunswick, New Jersey, United States

Brett Ashley Leeds
Rice University
Houston, Texas, United States

Lucas Leemann
Columbia University
New York City, New York, United States

Cristina Leston-Bandeira
University of Hull
Hull, United Kingdom

Thierry Leterre
Miami University John E. Dolibois Center
Luxembourg

David Levi-Faur
Hebrew University of Jerusalem
Jerusalem, Israel

Gang Lin
Shanghai Jiaotong University
Shanghai, China

Beate Littig
Institute for Advanced Studies
Vienna, Austria

Richard Little
University of Bristol
Bristol, United Kingdom

Steven Livingston
The George Washington University
Washington, D.C., United States

Francisco J. Llera
University of the Basque Country
Bilbao, Spain

Martin Lodge
London School of Economics and Political Science
London, United Kingdom

George A. Lopez
University of Notre Dame
Notre Dame, Indiana, United States

Hartmut Lüdtke
Philipps-Universität Marburg
Marburg, Germany

Samantha Luks
YouGov
Palo Alto, California, United States

Cora J. M. Maas (Deceased)
Utrecht University
Utrecht, Netherlands

Muiris MacCarthaigh
Institute of Public Administration
Dublin, Ireland

Jason A. MacDonald
West Virginia University
Morgantown, West Virginia, United States

Bogdan W. Mach
Polish Academy of Sciences
Warsaw, Poland

Mary Anne Madeira
University of Washington
Seattle, Washington, United States

Lars Magnusson
Uppsala University
Uppsala, Sweden

Martín A. Maldonado
Universidad Católica de Córdoba
Córdoba, Argentina

Paolo Mancini
Università degli Studi di Perugia
Perugia, Italy

Patrik Marier
Concordia University
Montréal, Québec, Canada

Radoslaw Markowski
Polish Academy of Sciences and Warsaw School of Social Sciences and Humanities
Warsaw, Poland

Emanuele Marotta
Ministry of Interior
Rome, Italy

Paolo Martelli
Università degli Studi di Milano
Milan, Italy

Danilo Martuccelli
University of Lille 3
Lille, France

Martin E. Marty
University of Chicago
Chicago, Illinois, United States

Alfio Mastropaolo
University of Turin
Turin, Italy

Pierre Mathiot
Institut d'Etudes Politiques de Lille
Lille, France

Irmina Matonyte
Institute for Social Research
Vilnius, Lithuania

Heinrich Matthee
Black Hall College
London, United Kingdom

Felicity Matthews
University of Exeter
Exeter, United Kingdom

Liborio Mattina
University of Trieste
Trieste, Italy

Gianpietro Mazzoleni
Università degli Studi di Milano
Milan, Italy

Rose McDermott
Brown University
Providence, Rhode Island, United States

Anthony McGann
University of Essex
Colchester, United Kingdom

Timothy J. McKeown
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina, United States

Philip McMichael
Cornell University
Ithaca, New York, United States

David E. McNabb
Pacific Lutheran University
Tacoma, Washington, United States

Mariana Medina
Washington University in St. Louis
St. Louis, Missouri, United States

Andreas Mehler
German Institute of Global and Area Studies
Hamburg, Germany

Andrei Y. Melville
Moscow State Institute of International Relations
Moscow, Russian Federation

Ines Mergel
Syracuse University
Syracuse, New York, United States

Wolfgang Merkel
Social Science Research Center Berlin (WZB)
Berlin, Germany

Peter H. Merkl
University of California, Santa Barbara
Santa Barbara, California, United States

Patrick A. Messerlin
Groupe d'Economie Mondiale at Sciences Po
Paris, France

Louis Meuleman
Nyenrode Business University
Breukelen, Netherlands

Jean-Baptiste Meyer
Institut de Recherche pour le Développement
Montpellier, France

Hugh Miall
University of Kent
Canterbury, United Kingdom

Katja Michalak
American University in Bulgaria
Blagoevgrad, Bulgaria

Carlos R. S. Milani
Universidade Federal da Bahia
Salvador, Bahia, Brazil

Gary J. Miller
Washington University in St. Louis
St. Louis, Missouri, United States

Helen V. Milner
Princeton University
Princeton, New Jersey, United States

José Álvaro Moisés
University of São Paulo
São Paulo, Brazil

Patrick Dibere Molutsi
Tertiary Education Council
Gaborone, Botswana

Paolo Moncagatta
Universitat Pompeu Fabra
Barcelona, Spain

James E. Monogan III
Washington University in St. Louis
St. Louis, Missouri, United States

Manuel Mora y Araujo
Universidad Torcuato Di Tella
Buenos Aires, Argentina

Laurence Morel
University of Lille II
Lille, France

Alejandro Moreno
Instituto Tecnológico Autónomo de México
Mexico City, Mexico

Anri Morimoto
International Christian University
Tokyo, Japan

Leonardo Morlino
LUISS Guido Carli
Rome, Italy

Ulrika Mörth
Stockholm University
Stockholm, Sweden

Rebecca B. Morton
New York University
New York City, New York, United States

Jonathon W. Moses
Norwegian University of Science & Technology
Trondheim, Norway

Karen Mossberger
University of Illinois at Chicago
Chicago, Illinois

Hans Mouritzen
Danish Institute for International Studies
Copenhagen, Denmark

Donald P. Moynihan
University of Wisconsin–Madison
Madison, Wisconsin, United States

Richard Mulgan
Australian National University
Canberra, Australia

Harald Müller
Peace Research Institute Frankfurt
Frankfurt, Germany

Ferdinand Müller-Rommel
Leuphana Universität Lüneburg
Lüneburg, Germany

Jordi Muñoz
Universitat Pompeu Fabra
Barcelona, Spain

Salvador Jimenez Murguia
Miyazaki International College
Miyazaki-gun, Miyazaki, Japan

David Mutimer
York University
Toronto, Ontario, Canada

Philip Rudolph Nel
University of Otago
Dunedin, New Zealand

Mark A. Neufeld
Trent University
Peterborough, Ontario, Canada

Iver B. Neumann
Norwegian Institute of International Affairs (NUPI)
Oslo, Norway

James L. Newell
University of Salford
Salford, United Kingdom

Steven Ney
Jacobs University Bremen
Bremen, Germany

Jairo Nicolau
Institute of Social and Political Studies, Rio de Janeiro State University
Rio de Janeiro, Brazil

Mark David Nieman
University of Iowa
Iowa City, Iowa, United States

Jolanda van der Noll
Bremen University
Bremen, Germany

Noburu Notomi
Keio University
Yokohama, Japan

Alfredo Nunzi
Europol, European Police Office
The Hague, Netherlands

Diana Z. O'Brien
Washington University in St. Louis
St. Louis, Missouri, United States

Anthony Oberschall
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina, United States

Claus Offe
Hertie School of Governance
Berlin, Germany

Santiago Olivella
Washington University in St. Louis
St. Louis, Missouri, United States

Andreas Osiander
Leipzig University
Leipzig, Germany

Abdulahi Osman
University of Georgia
Athens, Georgia, United States

Øyvind Østerud
University of Oslo
Oslo, Norway

Rachid Ouaissa
Philipps-Universität Marburg
Marburg, Germany

Joyce Outshoorn
Leiden University
Leiden, Netherlands

Henk Overbeek
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

Walter Ouma Oyugi
University of Nairobi
Nairobi, Kenya

Ergun Özbudun
Bilkent University
Anakara, Turkey

Leslie A. Pal
Carleton University
Ottawa, Ontario, Canada

Angelo Panebianco
University of Bologna
Bologna, Italy

Franz Urban Pappi
Mannheim University
Mannheim, Germany

Stéphane Paquin
École nationale d'administration publique
Montréal, Quebec, Canada

María del Carmen Pardo
El Colegio de México
Mexico City, Mexico

Charles F. Parker
Uppsala University
Uppsala, Sweden

Salvador Parrado
Universidad Nacional Educación a Distancia
Madrid, Spain

Misagh Parsa
Dartmouth College
Hanover, New Hampshire, United States

Gianfranco Pasquino
University of Bologna
Bologna, Italy

Werner J. Patzelt
Technical University of Dresden
Dresden, Germany

Ove K. Pedersen
Copenhagen Business School
Frederiksberg, Denmark

John Peeler
Bucknell University
Lewisburg, Pennsylvania, United States

Anthony W. Pereira
King's College London
Strand, United Kingdom

Sam Perlo-Freeman
Stockholm International Peace Research Institute
Solna, Sweden

B. Guy Peters
University of Pittsburgh
Pittsburgh, Pennsylvania, United States

Daniela Piana
University of Bologna
Bologna, Italy

Jon Pierre
University of Göteborg
Göteborg, Sweden

Michael Pinto-Duschinsky
Brunel University
Uxbridge, United Kingdom

Delphine Placidi
Institut d'Etudes Politiques de Toulouse (Sciences Po)
Toulouse, France

Thomas Plümper
University of Essex
Colchester, United Kingdom

Gianfranco Poggi
University of Trento
Trento, Italy

Keith T. Poole
University of California, San Diego
La Jolla, California, United States

Vincent Pouliot
McGill University
Montréal, Quebec, Canada

Michael Power
London School of Economics and Political Science
London, United Kingdom

Elisabeth Prügl
Florida International University
Miami, Florida, United States

Jerry Pubantz
University of North Carolina at Greensboro
Greensboro, North Carolina, United States

Michael Pye
Philipps-Universität Marburg
Marburg, Germany

Kevin M. Quinn
University of California, Berkeley
Berkeley, California, United States

Claudio M. Radaelli
University of Exeter
Exeter, United Kingdom

Charles C. Ragin
University of Arizona
Tucson, Arizona, United States

Frédéric Ramel
Université Paris 11 and Centre for Studies and Research of the Ecole Militaire
Paris, France

Luis Ramiro
University of Murcia
Murcia, Spain

Tiina Randma-Liiv
Tallinn University of Technology
Tallinn, Estonia

Brian Rathbun
University of Southern California
Los Angeles, California, United States

John Ravenhill
Australian National University
Canberra, Australia

James Lee Ray
Vanderbilt University
Nashville, Tennessee, United States

Philippe Raynaud
Université Panthéon-Assas (Paris 2)
Paris, France

Stefano Recchia
Columbia University
New York City, New York, United States

Marino Regini
Università degli Studi di Milano
Milan, Italy

Jeremy John Richardson
University of Oxford
Oxford, United Kingdom

Benoît Rihoux
Université Catholique de Louvain
Louvain-la-Neuve, Belgium

Jean-Jacques Roche
Université Panthéon-Assas
Paris, France

Bert A. Rockman
Purdue University
West Lafayette, Indiana, United States

Asbjørn Røiseland
Bodø University College
Bodø, Norway

Edeltraud Roller
Johannes Gutenberg University Mainz
Mainz, Germany

Paul G. Roness
University of Bergen
Bergen, Norway

Richard Rosecrance
Harvard University
Cambridge, Massachusetts, United States

William Rosenau
RAND Corporation
Arlington, Virginia, United States

Bo Rothstein
University of Göteborg
Göteborg, Sweden

Claude Rubinson
University of Arizona
Tucson, Arizona, United States

Dietrich Rueschemeyer
Brown University
Providence, Rhode Island, United States

Andrzej Rychard
Graduate School for Social Research
Warsaw, Poland

Ralf Rytlewski
Freie Universität Berlin
Berlin, Germany

Paul A. Sabatier
University of California, Davis
Davis, California, United States

Kurt Salentin
Bielefeld University
Bielefeld, Germany

Ari Salminen
University of Vaasa
Vaasa, Sweden

Nicholas Sambanis
Yale University
New Haven, Connecticut, United States

Cyrus Samii
Columbia University
New York City, New York, United States

Ulrich Sarcinelli
University Koblenz-Landau
Landau, Germany

Zeki Sarigil
Bilkent University
Bilkent, Ankara, Turkey

Anne E. Sartori
Northwestern University
Evanston, Illinois, United States

Sidsel Saugestad
University of Tromsø
Tromsø, Norway

Donald J. Savoie
Université de Moncton
Moncton, New Brunswick, Canada

Andreas Schedler
Centro de Investigación y Docencia Económicas (CIDE)
Mexico City, Mexico

Alexandra Scheele
University of Potsdam
Potsdam, Germany

Yves Schemeil
Institut d'Etudes Politiques de Grenoble
Grenoble, France

Theo Schiller
Philipps-Universität Marburg
Marburg, Germany

Antonio Schizzerotto
University of Trento
Trento, Italy

Peter Schmidt
National Research University Higher School of Economics (HSE)
Moscow, Russian Federation

Vivien A. Schmidt
Boston University
Boston, Massachusetts, United States

Philippe C. Schmitter
European University Institute
San Domenico di Fiesole, Italy

Ulrich Schneckener
University of Osnabrück
Osnabrück, Germany

Anne Schneider
Arizona State University
Tempe, Arizona, United States

Gerald Schneider
University of Konstanz
Konstanz, Germany

Saundra K. Schneider
Michigan State University
East Lansing, Michigan, United States

Helmar Schoene
University of Education Schwaebisch Gmuend
Schwaebisch Gmuend, Germany

Jan Aart Scholte
University of Warwick and London School of Economics
Coventry and London, United Kingdom

Lorna Schrefler
University of Exeter
Exeter, United Kingdom

Philip A. Schrodt
Pennsylvania State University
University Park, Pennsylvania, United States

Tobias Schulze-Cleven
University of Bamberg
Bamberg, Germany

Shirley V. Scott
University of New South Wales
Sydney, New South Wales, Australia

Jason Seawright
Northwestern University
Evanston, Illinois, United States

David Sebudubudu
University of Botswana
Gaborone, Botswana

Daniel-Louis Seiler
Aix-Marseille Université
Aix-en-Provence, France

Jasjeet S. Sekhon
University of California, Berkeley
Berkeley, California, United States

Sally C. Selden
Lynchburg College
Lynchburg, Virginia, United States

Mitchell A. Seligson
Vanderbilt University
Nashville, Tennessee, United States

Susan K. Sell
The George Washington University
Washington, D.C., United States

Martin Shapiro
University of California, Berkeley
Berkeley, California, United States

Doh Chull Shin
University of Missouri
Columbia, Missouri, United States

Alan Siaroff
University of Lethbridge
Lethbridge, Alberta, Canada

Perrine Simon-Nahum
Centre National de la Recherche Scientifique/
École des Hautes Études en Sciences Sociales
Paris, France

Luc Sindjoun
University of Yaoundé II
Yaoundé, Cameroon

Perri 6
Nottingham Trent University
Nottingham, United Kingdom

Chris Skelcher
University of Birmingham
Birmingham, United Kingdom

Gabriella Slomp
University of St. Andrews
St. Andrews, United Kingdom

Roger W. Smith
College of William and Mary
Williamsburg, Virginia, United States

Steven Rathgeb Smith
Georgetown University
Washington, D.C., United States

Fredrik Söderbaum
University of Göteborg
Göteborg, Sweden

Lourdes Sola
International Political Science Association and
University of São Paulo
São Paulo, Brazil

Eugenio Somaini
Università degli Studi di Parma
Parma, Italy

Albert Somit
Southern Illinois University
Carbondale, Illinois, United States

Eva Sørensen
Roskilde University
Roskilde, Denmark

Georg Sörensen
Aarhus University
Aarhus, Denmark

Bernardo Sorj
Federal University of Rio de Janeiro
Rio de Janeiro, Brazil

Arthur Spirling
Harvard University
Cambridge, Massachusetts, United States

Harvey Starr
University of South Carolina
Columbia, South Carolina, United States

Brent J. Steele
University of Kansas
Lawrence, Kansas, United States

Cindy Lee Steenekamp
Stellenbosch University
Matieland, South Africa

Janice Gross Stein
University of Toronto
Toronto, Ontario, Canada

Michael B. Stein
University of Toronto
Toronto, Ontario, Canada

Jennifer Sterling-Folker
University of Connecticut
Storrs, Connecticut, United States

Rudolf Stichweh
University of Lucerne
Lucerne, Switzerland

Edward H. Stiglitz
Stanford University
Stanford, California, United States

Laura Stoker
University of California, Berkeley
Berkeley, California, United States

Tara W. Stricko
Kennesaw State University
Kennesaw, Georgia, United States

Jörg Strübing
Universität Tübingen
Tübingen, Germany

Patrycja J. Suwaj
Bialystok University
Bialystok, Poland

Hiroki Takeuchi
SMU
Dallas, Texas, United States

Finn Tarp
University of Copenhagen
Copenhagen, Denmark

Daniel Tarschys
University of Stockholm
Stockholm, Sweden

Markku Temmes
University of Helsinki
Helsinki, Finland

Jan Teorell
Lund University
Lund, Sweden

Benno Teschke
University of Sussex
Brighton, United Kingdom

Henry Teune
University of Pennsylvania
Philadelphia, Pennsylvania, United States

Mark Thatcher
London School of Economics and Political Science
London, United Kingdom

Cameron G. Thies
University of Iowa
Iowa City, Iowa, United States

Victor Thiessen
Dalhousie University
Halifax, Nova Scotia, Canada

Clive S. Thomas
University of Alaska Southeast
Juneau, Alaska, United States

Craig W. Thomas
University of Washington
Seattle, Washington, United States

Mark R. Thompson
Friedrich-Alexander University Erlangen-Nuremberg
Erlangen, Germany

Bassam Tibi
University of Goettingen
Goettingen, Germany

Guido Tiemann
Institute for Advanced Studies, Vienna
Vienna, Austria

Keith Topper
University of California, Irvine
Irvine, California, United States

Mariano Torcal
Universitat Pompeu Fabra
Barcelona, Spain

Jacob Torfing
Roskilde University
Roskilde, Denmark

Shawn Treier
University of Minnesota
Minneapolis, Minnesota, United States

Carlo Trigilia
University of Florence
Florence, Italy

Vera E. Troeger
University of Essex
Colchester, United Kingdom

Peter F. Trumbore
Oakland University
Rochester, Michigan, United States

Ilter Turan
Istanbul Bilgi University
Istanbul, Turkey

Colin Tyler
University of Hull
Hull, United Kingdom

Arild Underdal
University of Oslo
Oslo, Norway

Camil Ungureanu
Universitat Pompeu Fabra
Barcelona, Spain

Bjørn Olav Utvik
University of Oslo
Oslo, Norway

Jan W. van Deth
University of Mannheim
Mannheim, Germany

Carolien Van Ham
European University Institute
San Domenico di Fiesole, Italy

Frans K. M. van Nispen
Erasmus University Rotterdam
Rotterdam, Netherlands

Dana R. Vashdi
University of Haifa
Haifa, Israel

Pascal Vennesson
University Paris II, European University Institute
San Domenico di Fiesole, Italy

Bram Verschuere
University College Ghent & Ghent University
Ghent, Belgium

Wytske Versteeg
Universiteit van Amsterdam
Amsterdam, Netherlands

Luca Verzichelli
University of Siena
Siena, Italy

Eran Vigoda-Gadot
University of Haifa
Haifa, Israel

Ulrich von Alemann
University of Düsseldorf
Düsseldorf, Germany

Klaus von Beyme
University of Heidelberg
Heidelberg, Germany

Marisa von Bülow
Universidade de Brasília
Brasília – DF, Brazil

Astrid von Busekist
Sciences Po
Paris, France

Patrick von Maravic
Zeppelin University
Friedrichshafen, Germany

Thomas von Winter
University of Potsdam
Potsdam, Germany

Claudius Wagemann
Istituto Italiano di Scienze Umane
Florence, Italy

Pieter Wagenaar
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

Uwe Wagschal
University of Heidelberg
Heidelberg, Germany

Lee Demetrius Walker
University of South Carolina
Columbia, South Carolina, United States

Christopher M. Weible
University of Colorado Denver
Denver, Colorado, United States

Barry R. Weingast
Stanford University
Stanford, California, United States

Jutta Weldes
University of Bristol
Bristol, United Kingdom

Jennifer M. Welsh
University of Oxford
Oxford, United Kingdom

Christian Welzel
Jacobs University Bremen
Bremen, Germany

Bettina Westle
Philipps-Universität Marburg
Marburg, Germany

Guy D. Whitten
Texas A & M University
College Station, Texas, United States

Jerzy J. Wiatr
European School of Law and Administration
Warsaw, Poland

Catherine Wihtol de Wenden
CERI-Sciences Po/CNRS
Paris, France

Ulrich Willems
University of Münster
Münster, Germany

Peter Willetts
City University, London
London, United Kingdom

Kenneth C. Williams
Michigan State University
East Lansing, Michigan, United States

Søren C. Winter
Danish National Centre for Social Research
Copenhagen, Denmark

Anders Wivel
University of Copenhagen
Copenhagen, Denmark

Edmund Wnuk-Lipinski
Institute of Political Studies, University of Warsaw
Warsaw, Poland

Hellmut Wollmann
Humboldt-Universität Berlin
Berlin, Germany

Byungwon Woo
Oakland University
Rochester, Michigan, United States

James Woods
Administrative Office of the U.S. Courts
Washington, D.C., United States

Mikhaïl Dorel Xifaras
Sciences Po
Paris, France

Kutsal Yesilkagit
Utrecht University
Utrecht, Netherlands

Brigitte Young
University of Münster
Münster, Germany

I. William Zartman
Johns Hopkins University
Baltimore, Maryland, United States

Richard O. Zerbe Jr.
University of Washington
Seattle, Washington, United States

Jan Zielonka
University of Oxford
Oxford, United Kingdom

Ekkart Zimmermann
Dresden University of Technology
Dresden, Germany

Andrei Zinovyev
Institut Curie
Paris, France

Anthony R. Zito
Newcastle University
Newcastle upon Tyne, United Kingdom

Preface

Politics has always been a key dimension of human life. Politics affects everyone, just as the weather does, and since the time of the Greek classics, humankind has attempted to come to grips with its empirical and normative implications. Politics and political science have changed over time and have been even more deeply transformed during the past decades in every part of an ever-smaller world. Empirical research and theoretical reflections on politics and its multiple connections with all other aspects of human life have developed enormously during the second half of the past century and now cover virtually all parts of the world, with their growing interdependence. These studies concern, for example, basic issues such as war and peace, prosperity, welfare, and a sustainable environment, as well as issues of freedom, justice, gender, and democracy under changing cultural perspectives.

At this point, one feels the need for an exhaustive overview of the empirical findings and the reflections on politics. The *International Encyclopedia of Political Science* caters to that need in many different kinds of readers, including undergraduate or graduate students who like to be informed effectively and quickly on their field of study; scholars who seek information on the relevant empirical findings in their area of specialization or in related fields; and lay readers, who may not have a formal background in political science but are attracted to, and interested in, politics. Thus, the encyclopedia can serve both the diffusion of knowledge about politics and the further development of the field.

The encyclopedia is the result of collaborative work by leading international scholars from all over the world. This team ensured the widest possible coverage of key areas, both regionally and globally. In this sense, this is the first truly international encyclopedia of its kind. It aims to give a comprehensive picture of all aspects of political life, recognizing the theoretical and cultural pluralism of the approaches and including findings from all parts of the world. The eight volumes of the encyclopedia cover every field of politics, from political theory and methodology to political sociology, comparative politics, public policies, and international relations.

The entries are arranged in alphabetical order, and a list of entries by subject area appears in the front of each volume. The entries are organized in three major categories, according to their substantive relevance, of roughly 1,800, 3,600, and 8,000 words. In addition, longer entries of 12,000 words cover the major subdisciplines and the state of the art in each field. The encyclopedia contains a detailed index as well as extensive bibliographical references. Thus, it can provide an essential and authoritative guide to the state of political science at the beginning of the 21st century for decades to come. It will be an invaluable resource for a global readership, including researchers, students, and policymakers.

The editors wish to acknowledge the invaluable assistance and advice of many colleagues and friends, including the members of the Executive Committee of the International Political Science Association (IPSA) and its presidents during this period, Max Kaase and Lourdes Sola. Bertrand Badie wishes to thank Sciences Po, Paris, and particularly Delphine Placidi, who assisted him in the first steps of the work. Dirk Berg-Schlosser gladly acknowledges the assistance from Diana Mai and Joris-Johann Krapp at the Institute of Political Science, Philipps-University, Marburg, Germany, in the initial phases of the project. Leonardo Morlino would

like to warmly thank Dr. Claudius Wagemann for his invaluable help in the starting phase of the encyclopedia.

Bertrand Badie
Professor of International Relations,
Paris Institute of Political Studies (Sciences Po)
Paris, France

Dirk Berg-Schlosser
Emeritus Professor of Political Science,
University of Marburg
Marburg, Germany

Leonardo Morlino
Professor of Political Science, LUISS
Rome, Italy

Introduction to Political Science

> *Cet ouvrage produira sûrement avec le temps une révolution dans les esprits, et j'espère que les tyrans, les oppresseurs, les fanatiques et les intolérants n'y gagneront pas. Nous aurons servi l'humanité.* [Over time, this work will surely produce an intellectual revolution, and I hope that the tyrants, oppressors, fanatics, and those without tolerance will not benefit from it. We will have served humanity.]
>
> —Letter from Diderot to Sophie Volland, September 26, 1762

Political science, paradoxically, is both a very old and relatively recent discipline. Its origins go back to antiquity in classic European or Asian thought as far as recorded history goes. As an independent and respected academic field, however, it came into being in most countries only after World War II. This is due in part to the fact that its subject matter had been addressed by neighboring disciplines like philosophy, history, and public law, but also because it requires, more than others, a "breathing space" of freedom of thought and expression that is not voluntarily granted by most authoritarian regimes and that has developed worldwide only with decolonization and recent waves of democratization.

This introduction surveys the discipline of political science, beginning with an analysis of politics itself. Two important definitions of politics are discussed in detail here. First, politics can be viewed as a means, either for maintaining social harmony or for achieving a supreme good, as in some religious conceptions of the state. Second, politics has been understood as an instrument by which a central institution or a legitimate authority exercises power. The next section of the introduction examines the epistemological foundations of the discipline of political science. The third section traces the growth of political science during the 20th century and examines its evolving relationships to other social sciences and to other fields, most notably law and philosophy. The final section of the introduction describes more recent developments and perspectives in political science, as it has become a dynamic discipline with its own identity and with participation from political scientists around the globe.

Politics

Politics as a Special Sphere

The concept of politics carries different meanings. It can be considered to be an *art* (scholars are "studying politics"); an *activity* (one can "play politics" in one's office, in one's club, even in one's family); a *profession* (some "go into politics"); or a *function* ("local politics," "national politics"). Most political scientists consider the first meanings as derived from the last one, even as metaphoric, while some others have a wider and more abstract conception that is broader than that of function. Politics also has been understood as both to polity, which refers to an organization (a state, a regime and its constitution) and to policy, which refers to a system of political decisions and specific subfields (like health, education, foreign relations) through organizations act to carry out their functions. In a comprehensive perspective, politics is thus considered to be linked to a function, a system, an action, and a behavior.

All of them are, however, connected to a special dimension of the history of humankind. Even if some scholars object that some societies ignored politics (Clastres, 1975), most anthropologists consider politics as a constant of the human condition. In the first part of this entry, this constant will be grasped in its various definitions, and then it will be inserted into the general social order. The following sections examine definitions of this constant as a function and as an instrument.

Politics as a Function

Politics can be conceived as a contribution to the social adventure, as a function of the social order, or, quite differently, as a distinctive instrument, a special way of action. The first approach is more classic and deeply rooted in the various philosophical traditions that were elaborated around the world, while the second one is modern, related to the rise of positivist theory in the social sciences. In this section, we consider two ways in which politics can be seen as a function.

Politics as a Function: Promoting Social Harmony or Supreme Good?

Many philosophers have located politics in the art of coexistence. If human beings are selfish by nature, as they are often conceived, but must live and grow up together, to create peaceful coexistence is obviously one of the main functions of the polis. As such, politics should be considered as the permanent invention of the polis (city), as the construction of each social unit that aims to keep people together on a permanent basis. This point was already made by Plato, who considered politics as the art of organizing social harmony. We also find it in other traditions. Islam conceives of politics as a weakly differentiated function that aims to overcome tribal fragmentation through the principle of unity (*tawhid*). As such, *tawhid* will be achieved through the absolute Unity of God, and so politics cannot be entirely accomplished without religion: Politics cannot be conceived as a differentiated structure, but it is obviously a social function. The same principle can be found in Hindu writings. The *Arthasastra* (3rd century BC) and the *Manusmriti* (2nd century BC) were written during periods of decay and so-called evil, which implied starvation, violence, and chaos, while the *Mahabarata* covered a much longer period (from 1st millennium BCE to the 5th century BCE). Here, politics is presented as an absolute requisite for keeping peace and order, even if the tradition oscillates between divine invention in politics (*Manusmriti*) and its contractual origin (*Mahabarata*). In its turn, the Buddhist vision stresses the entropy of the world that leads to inequality, sexual division, property, and thus to conflicts and lack of safety. The orientation of this reform of Hinduism prompts human beings to choose a king as a guarantor of the social order. Similarly, Confucius pointed out that men need a ruler for preventing disorder, disturbance, and confusion.

This first function, promoting social harmony, obviously shaped political philosophy to a large extent up to the present day. Social contract theory clearly emanated from this postulate, in the Islamic tradition (*mithaq, bay'a*) as well as during the European Enlightenment. The *Arthasastra* anticipated Thomas Hobbes's vision of the state of nature, when it described the lack of politics as resulting in evil and vices, or when it mentioned the fable of the big fish that will eat the small one (*Matsyanyaya*). The functional dimension of politics, as the art of coexistence and "maintaining harmony," can, therefore, be considered as really transcultural and common to the humanity's different histories. Here, we can probably locate the roots of a pluralist vision of politics, as this first definition paves the way for a plural conception of the city (polis) where people do not necessarily share the same interests, the same beliefs, or the same ethnic characteristics. Quite the opposite, in this perspective, diversity is the real raison d'être of politics.

However, politics also claims a second function, which is more demanding and sophisticated. Some philosophers and thinkers are going further, beyond the invention of the city, pointing to another purpose: Politics is supposed to lead to the path of righteousness, to promote virtue, and to enable humans to achieve the Supreme Good. Thus, Aristotle conceived of politics as referring to welfare and virtue. The city must be constructed as the *good city*: Political science is elevated then to something much more demanding, namely, the "science of the good politics" or the "science of good government." This vision can be found in Islam through the commitment to divine law (*sharia*); when taken to its extreme, this conception even becomes a way of challenging power holders and leads to a political inversion in which protest is a more important political activity even than governing. In the same way, the *Arthasastra* describes politics as promoting peace and prosperity, while Buddhism produces an ethic of human behavior. Confucianism also gives the central role to the virtue that humans naturally possesses, but that is achieved through the ruler and his norms.

These two functions of politics convey the two faces of political theory, one of which is positivist, the other normative. If politics is only the science of the city, it is first of all a *behavioralist* science. If it is the science of the Supreme Good, its *normative* orientation is dominant. This tension has partly been overcome by an instrumentalist approach to politics, according to which the distinguishing nature of politics has to be found in the *instruments* used for running the city instead of its ultimate goals.

Politics as an Instrument: The Science of Power?

Power—in Max Weber's sense as the ability to achieve your interests even against someone else's will, that is, as coercion—is understood here as the first instrument through which politics operates. All major thinkers claim that the city cannot exist without power, no matter how it is structured. There is thus a long tradition of connecting power and politics, in which political science is assigned to study how power is formed, structured, and shared (Lasswell & Kaplan, 1950). However, the *science of politics* and a *science of power* are not synonymous: Is every kind of power necessarily political, for example, in a firm or a club? In a broader sense, some authors state that politics can be played in an office or inside a family, but this expression is merely a metaphor when conceived in a micro-social order. Conversely, if we consider power as the essence of politics, we have to opt for a wide definition of power that includes ideology, social control, and even social structures, and thus deviates from a vision of power as purely coercive.

For this reason, politics is commonly defined as a specific kind of power; either it is held by a central institution, such as a state, a government, a ruling class, or it is used by a power holder who is considered to be legitimate. The first perspective approaches politics as the science of the state and implies that traditional and weakly institutionalized societies lack the centralized power that is necessary for politics to exist. The second promotes the concept of legitimate authority as the real essence of politics and suggests that in this sense politics is more evident in democracies than in than authoritarian or totalitarian systems.

Some visions also link power and norms. Pushing the Aristotelian definition further, they define politics as the set of norms that lead to the City of Good—either the City of God or the City of the Philosopher. This radical conception is to be found in several traditions: that of Augustine, rather than Thomas Aquinas, in the Christian culture; Puritanism with the Reformation and Calvinism; radical Islam in the perspective of Ibn Taimyya. This conception, however, runs the risk of drifting into totalitarianism or at the very least depreciating the political debate, as it make any kind of political choice impossible.

The connection with *territory* is also frequently used as an instrumental approach to politics. The Greek tradition paved the way when Aristotle stressed the difference between politics, *ethnos,* and *oîkos* (household). If politics is conceived as the coexistence of diverse peoples, it denies that there is such a thing as "natural territory" and implies a socially constructed territory as its required arena. That is why Max Weber makes territory a key element of his definition of politics. For Weber, a community has a political quality only if its rules are granted inside a given territory: this territorial invention hardly fits nomadic societies or even a number of traditional ones (Evans-Pritchard, 1940). But even if it gets close to a state vision of politics, it emphasizes the role of pluralism and diversity inside the political order. In a more extensive conception of the spatial dimension of politics, public and private spheres are opposed: the former is seen as the natural background of political debate, while the latter is conceived as a resistance against political power and its penetrations (Habermas, 1975). We also find again the possible opposition between religious and secular spheres (other-worldly/this-worldly), and even the disenchantment with the world as one of the possible sources of politics.

After going through all these definitions, the criterion of social coexistence seems to be the most extensive one, and probably the least questionable. If politics is everywhere around the world considered more or less as managing social harmony, it can clearly be conceived as the opposite of some other classical spheres of social action (politics vs. social life, military, administration, etc.). As such, it is part of the general social arena, as an ordinary social fact, but a very specific one.

Politics in the Social Division of Labor

Here we face a contradiction that is shaping a serious debate among political scientists. If the conception of politics as an ordinary social fact tends to prevail, political science merges with political sociology (see below). In the opposite version, the latter would be defined as a part of political science, sometimes with ambiguous borderlines. The vagueness and the mobility of the borderline stem from different factors: the diversity of the great theories in the social sciences, which do not reflect the same visions of politics and which are torn between power and integration; the historical and cultural background of politics, which is shaping different kinds of lineages; and the present impact of globalization, which is probably fueling a new definition of politics that is increasingly detached from concepts of ethnicity and territoriality.

Two Traditions: Power Versus Integration

Max Weber is obviously considered as a "founding father" by both sociologists and political scientists. He clearly promoted a political vision of sociology when he developed his two major concepts of *Macht* (power as coercion) and *Herrschaft* (power as authority). Both of them can be found in the very first steps of his sociology where he defines power as the ability of one actor in a social relationship to modify the behavior of another, through pressure, force, or other forms of domination. From a Hobbesian perspective, power plays the major role in structuring social relationships, while the social actors strive to give meaning to this asymmetrical relationship in order to make it just and acceptable, thus establishing the legitimacy of those with power.

In modern society, the state plays an important role and politics has an exceptional status, as it is theoretically conceived as the main basis of social order. This conception is also strongly rooted in the Marxist vision, where the state is considered the instrument by which the ruling class maintains its domination, as the bourgeoisie does in the capitalist mode of production. Carl Schmitt also starts from a Weberian presupposition in linking politics to enmity. By contrast, in a Durkheimian vision, integration is substituted for power as the key concept. Politics is conceived neither as an instrument of domination nor as a way of producing social order; rather, it is a function by which the social system is performing its integration. Obviously, this function implies institutions and then a political sphere, including state and government, but it is considerably more diffuse and appears to be produced by the social community and its collective consciousness. From this perspective, a political society is made up of social groups coming together under the same authority. Such an authority derives from the social community and the collective consciousness; it is constituted by rules, norms, and collective beliefs, which are assimilated through socialization processes. As such, politics is closely related to social integration and is supposed to strengthen it further. That is why there is a strong correlation between a growing division of labor, from the increasing political functions, and their differentiation from the social structures. "The greater the development of society, the greater the development of the state" (Durkheim, 1975, 3, p. 170). Durkheim contrasted "mechanical solidarity" arising because of perceived similarities among people (e.g., in work or education) from "organic solidarity" arising when people are doing different things but see themselves as part of an interdependent web of cooperative associations. This Durkheimian vision is to be found later in the functionalist and systemic concepts of political science as elaborated by Talcott Parsons, David Easton, Gabriel Almond and others, but also in the socio-historical traditions, which attempted to link the invention of politics to the sociology of social changes, as in the work of Charles Tilly or Stein Rokkan. It is also congruent with the social psychological paradigm, which tries to capture politics through its social roots, such as socialization, mobilization, and behavioral analysis.

By contrast, state and power are the real sources of a Weberian political science. Politics is no longer a function of the division of labor, but has definitely its own determinants. Quite the opposite, social history is considering the transformations in the mode of government and more precisely the mode of domination. Power is thus conceived as an explanatory variable of the transformation of societies and political orders. Such a vision is common among those approaches of political science that are centered on power politics, the role of the state, or the nature of political regimes or that are focusing on political institutions and the conditions of their legitimization.

The Diversified Lineages of Politics

Politics is thus approached in different ways, but is also intrinsically plural. During the 1970s, when globalization began to shape the world and when decolonization was completed, both history and anthropology incorporated the perspective of politics with respect to plurality. This perspective on plurality also challenged the mono-dimensional vision that had been promoted by developmentalism a decade earlier. In anthropology, Geertz (1973) pointed out that politics covered several meanings that are changing along historical lines and according to specific cultures. These meanings are socially constructed as human actors encounter different kinds of events, challenges, or goals and as they are rooted or embedded in different sorts of economic and social structures. Politics is understood as achieving the will of God and his law in Islam, while it aims to manage the human city in this world according to the Roman Christian culture. The first conception was fueled by Muhammad's *hijra* when the Prophet left Mecca because of opposition to his teaching and went to Medina to build up the City of God. The second conception was shaped by the Roman experience of religion, which survived during the centuries of the Empire and had again to survive when the latter collapsed during the fifth century. In this dramatic contrast, politics does not cover the same meaning, as it is differentiated from the public sphere and oriented toward individuals in the Roman tradition, while it was more globally constructed in the Muslim tradition. In both cultures, this diversification continued; as Geertz mentions, politics does not have the same meaning in Indonesia and in Morocco, two Muslim societies that experienced greatly different histories.

For that reason, politics can be properly defined only when the definition includes the meaning that the social actors usually give to it. This cultural background implies a huge empirical investigation, which is all the more difficult since the observer tends to view things through his or her own concepts, which are obviously culturally oriented. The risk, therefore, is high to consider as universal a cultural vision of politics, which shapes the paradigm of empirical political science. Translations can be particularly misleading and even fanciful. For instance, the Arabic word *dawla* is often translated as "state," whereas their meanings are hardly equivalent. It is quite impossible to convey, through translation, the deep cultural gap that really implies two competing visions of politics. The only way of going ahead is to deepen the anthropological and the linguistic investigations in order to identify distinctive features of each conceptualization of politics, following the "thick description method" recommended by Geertz. But is there any end to this "individualization" of politics? To be operative, research must postulate a minimal universality of its own concepts and contain the risks of "culturalism": It has to keep the connection with history and anthropology while remaining in a universalist framework.

A Politics of Globalization

This dilemma is revived and even stimulated by the globalization of the world. In the new global order, politics is no longer limited or contained by the territoriality principle. It becomes reinvented beyond the classical coexistence of sovereign cities. Politics cannot be conceived as a simple addition of social contracts, as it was in the Westphalian paradigm. This challenge is first posed to the "realist" theory of international relations, questioning the absolute opposition between "inside" and "outside," or "domestic politics" and "international politics." The latter is no longer confined to the dialogue of sovereigns and has destroyed the traditional categories and criteria of politics. After all, is there a "global covenant," as Robert Jackson (2000) argues, that totally reshapes the construction of politics?

The hypothesis that competition among nation-states can be understood as parallel to that within nation-states supported the extension of the concept of politics to the international sphere. The idea of power politics was projected into the international arena in order to stress that international politics referred to the classical grammar: States, like political actors, were competing according to their own interests and were primarily concerned with their ability to dominate other states, or, at least, to contain the power of the others. Morgenthau (1948) defined *international politics* as the "struggle for power," *power* as "the control over the minds and actions of other men," and *political power* as "the mutual relations of control among the holders of public authority and between the latter and the people at large" (p. 27).

Although this conception is clearly rooted in a Weberian approach to politics, it does not belong only to the past. But it neither covers nor exhausts all the political issues at stake in the new configuration of the international arena. First of all, as sovereignty is fading, the proliferation of transnational actors no longer restricts international politics to a juxtaposition of territorial nation-states. Second, power and coercion are losing their efficiency as influence and social relationships are getting more and more performance oriented. Third, globalization and the growing international social community are shaping common goods, creating a kind of community of humankind; human beings are then creating "a political dialogue that can bridge their differences . . . without having to suppress them or obliterate them" (Jackson, 2000, p. 16). We are here rediscovering Aristotle when he claimed that men need each other for their own survival.

Nevertheless, no one would assert as yet the complete achievement of an international society or an international community. International politics remains an unstable combination of references to power politics and to international social integration: It then confronts the vision of politics as coexistence among diversity. It goes back to the idea of harmony, but without a completed contract, to the hypothesis of a global city without a central government, to the assertion of common norms without binding measures. This combination is at the core of the English School of international relations that refers to the "anarchical society." But it is also close to the French vision of an international solidarity. In the end, politics gets closer and closer to a functional vision of managing social diversity in order to make it compatible with the need for survival.

Epistemological Building Blocks of Political Science

Some major building blocks of political science can be identified that help characterize some common elements in existing approaches, but also, and perhaps more important, enable us to locate these positions and their *differences* more precisely with regard to the major epistemological foundations. The first of these building blocks concerns the *multidimensionality* of our subject matter; the second, its *plastic* and malleable character and the resulting *self-referential* problems; the third refers to a *systems* perspective of politics; and the fourth to the *linkages* between different levels (micro-, meso-, macro-) of political (and more generally social) analysis.

Multi-Dimensionality

The most basic distinctions of our discipline, which distinguish it in some important respects from the "natural" sciences, concern the dimensions of its subject matter. As in nature, there are certain hard "objects" such as political institutions and social structures, which can be identified and which are "tangible" and observable in certain ways. In addition, however, there is a "subjective" dimension in which such objects are perceived by individuals and groups and translated into concrete actions. Such perceptions themselves are shaped by a number of psychological, social, or other factors. This distinction is commonly accepted and runs through the history of philosophy from antiquity to the present day and concerns all sciences of humankind, including medicine. There, distinctions between body and mind (or consciousness) and the subsequent divisions into subdisciplines such as anatomy and psychology are commonplace. Similarly, the fact that there are possible interactions between these dimensions is well accepted, even though in medicine some of these psychosomatic relationships are still not well researched. The third dimension, the "normative" one that concerns ethical judgments of "good" or "bad" actions and behavior, is more problematic. In medicine, again, some ethical norms have been generally accepted since the time of Hippocrates, but debates continue about, for example, when exactly human life begins or ends, and what the respective theological or philosophical justifications are for such positions. In philosophy, this "three-dimensionality" of human existence has also been elaborated by Immanuel Kant, for example, in his *Critique of Pure Reason* (1787/1956, p. 748 ff.).

A graphical representation of these dimensions can be rendered in the following Figure 1 (where the dotted line represents a "holistic" position):

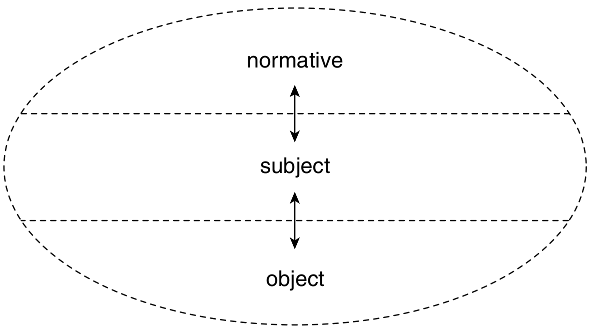


Figure 1 Dimensions of Human Existence

Source: Dirk Berg-Schlosser.

The crux of the matter really concerns problems of distinguishing such dimensions and their interactions not only analytically but also in actual practice, and controversies about normative, ontologically based justifications and their respective epistemological and methodological consequences persist. Here, we cannot go into these debates in any detail, but we find it useful to locate the major emphases of the current meta-theoretical positions in political science with the help of such distinctions. Thus, the major ontological approaches have their basis in the normative dimension ranging from Plato to Eric Voegelin or Leo Strauss, but also concern attempts in linguistic analysis (e.g., Lorenzen, 1978), or communications theory (Habermas, 1981). This also applies to non-Western traditions such as Confucian (Shin, 1999), Indian (Madan, 1992), or sub-Saharan African (Mbiti, 1969) ones.

Sharply opposed to such foundations of political theory are critical-dialectical or historical-materialist positions in the tradition of Karl Marx and his followers. There, the object dimension of the modes of production and re-production of human existence is the basic one from which the others are derived. Thus, the objective social existence determines the subjective consciousness and the political and normative superstructures. On this position, the teleological theory of history of Marx and his followers from the early beginning until the classless society and its peaceful end is based as well.

The third major meta-theoretical position, a behavioral or behavioralist one, takes the subjective dimension as its starting point, expressing a position of methodological or phenomenological individualism (for the use of these terms, see Goodin & Tilly, 2006, p. 10ff.). Then subjective perceptions and subsequent actions of human beings are what really matters. These shape social and political life. This position has been most influential in election studies, for example, but also concerning some aspects of political culture research. In a somewhat broader perception, both subjective and objective dimensions and their interactions are considered by empirical-analytical approaches, but, from a positivistic point of view, no normative judgments can be made on this basis. Long-lasting controversies concerning this position go back to Max Weber and his followers but are also reflected in more recent debates between Karl Popper and Jürgen Habermas, for example (see Adorno et al., 1969).

These basic meta-theoretical positions and their variations remain incompatible. Similarly, whether these dimensions can in actual fact be separated or, by necessity, always go together from a holistic perspective remains controversial. The latter position, in contrast to Kant, is, for example, represented by G. W. F. Hegel, but also by Marx and some of his followers (e.g., Lukács, 1967). In the same way, epistemological positions based on religion, including Buddhism and Confucianism, perceive these dimensions in a holistic manner. From a more pragmatic perspective, it seems that the fundamentalist debates about such matters have subsided in the last few decades and most political or social scientists just agree to disagree about such basic ontological or religious positions and their respective justifications. Nevertheless, Figure 1 may help better locate such positions and to put some conceptual order into these controversies.

The Plastic Matter of Political Science

As a result of the change from Newtonian physics with its deterministic relationships to quantum theory and probabilistic relations in nuclear physics, Popper (1972) has coined the metaphor of "clouds and clocks." Clocks represent Isaac Newton's deterministic world, as in astronomy, for example, where the movements of stars and planets or the next solar eclipse can be predicted (or retro-dicted) with clocklike precision. Clouds, by contrast, constitute a very elusive substance, the structures and regularities of which cannot easily be grasped over a somewhat longer period even today by the most advanced computers of meteorologists and their satellite-based data. Between these two extremes, however, which should be perceived as the opposite poles of a continuum rather than mutually exclusive positions, there is a plastic matter that is malleable in the course of time and that is neither perfectly determined nor subject to pure chance.

In an important essay, Gabriel Almond and Stephen Genco (1977) have transferred this concept to the social sciences and politics. They state that

> the implication of these complexities of human and social reality is that the explanatory strategy of the hard sciences has only a limited application to the social sciences. . . . Thus, a simple search for regularities and lawful relationships among variables—a strategy that has led to tremendous success in the physical sciences—will not explain social outcomes, but only some of the *conditions* affecting those outcomes. (p. 493, emphasis added)

The deductive subsumption of individual events under "covering laws" in Carl Hempel's (1965) sense, according to which claims about individual events can be derived deductively from premises that include a scientific law, thus is not possible for the most part. In addition, factors of human *choice* and action plus, possibly, some elements of pure chance in certain conjunctures also have to be considered.

As a consequence, we have to be more modest in our claims about the precision of causal relationships, the generalizability of regularities, and the universality of theories. At best, therefore, only theories located more precisely in time and space—what Robert K. Merton called "medium-range theories"—seem to be possible for most practical purposes. Such a view also corresponds with a position already expressed by Aristotle, who located politics in an intermediate sphere between the necessary, where strict science can be applied, and the realm of pure chance, which is not accessible for scientific explanations.

Such distinctions are illustrated in Figure 2:

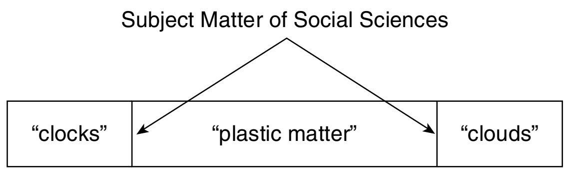


Figure 2 Degree of Determination of Theories

Source: Dirk Berg-Schlosser.

Again, the full implications of such a perspective cannot be discussed here, but this figure should be helpful, once more, to locate some of the "harder" and some of the "softer" approaches in our discipline along this spectrum. On the whole, we would agree with Almond and Genco's conclusion that

> the essence of political science . . . is the analysis of choice in the context of constraints. That would place the search for regularities, the search for solutions to problems, and the evaluation of these problems on the same level. They would all be parts of a common effort to confront man's political fate with rigor, with the necessary objectivity, and with an inescapable sense of identification with the subject matter which the political scientist studies. (p. 522)

The last point also leads to the next *differentia specifica* of the social sciences as compared to the naturalist sciences and their distinct epistemology.

Self-Referential Aspects

This sense of identification also can be seen in different ways. First of all, it means that as human and social beings we are inevitably part of the subject matter we are studying. Even if we attempt to detach ourselves as much as possible from the object under consideration some subjective influences on our perception remain. These can be analyzed by psychology, the sociology of knowledge to discern our (conscious or unconscious) "interests" in such matters, and so on, but some individual "coloring" of our lenses seems inevitable. Therefore, a certain "hermeneutic circle," which should be made conscious and explicit in the interactions with others, remains (Moses & Knutsen, 2007, Chapter 7).

However, this limitation can, again in contrast to naturalist perceptions of science, be turned to one's advantage. As human beings we can empathize with each other and can intersubjectively, if not objectively, understand and interpret the meaning of each other's thoughts and actions. This is even more the case when we are trained as social scientists in a common methodology and scientific language. This latter point also distinguishes the perception, level of information, and theoretical interpretation of a political scientist from the "man (or woman) in the street" talking politics, in the same way that a meteorologist has a different knowledge of what is happening in the atmosphere compared to the daily small talk about the weather. Nevertheless, such inevitable subjectivity, which is also historically and culturally conditioned, opens the way to more pluralist interpretations and meanings. Constructivist approaches, as contrasted to naturalist ones, can dig deeper in certain ways into this subjectivity and the plurality of meanings (cf., e.g., Foucault, 1970).

Two more points concerning our identification with the subject matter and our self-referential position within it must be mentioned. Being part of the substance, we can also, consciously or unconsciously, act upon it. Thus, self-fulfilling or self-defeating prophecies become possible as feedbacks between the interpretation or even just personal opinion of an important actor or social scientist whose authority in a certain sphere has become acknowledged in the matter he is dealing with. This frequently occurs when some "analysts" give their opinion on probable developments of the stock exchange or currency rates and many people follow suit. This also applies to electoral predictions with respective bandwagon and underdog effects.

Finally, being part of our world and being able, to some extent, to act on it, also raises the question of social and political responsibility. This brings us back to the normative side of politics with which we inevitably have to deal, self-consciously and being aware of possible consequences. In this respect, too, a recent constructivist turn in the theory of international politics, in a somewhat more specific sense of the term, has led to the broader discussion and possible acceptance of more universal norms.

A Systems Perspective

Within this multidimensional, malleable, and dynamic universe more specific *political* elements can be identified. One difficulty in this respect, again, lies in the contrasting meta-theoretical positions and their perspectives on politics (see the first section above). In a more abstract way, politics can also be conceived as the regulating mechanism in large-scale modern societies. Easton (1965) thus defines politics as "the authoritative allocation of values" in society and the forces shaping these procedures. In this process, different elements interact in a systemic way regulating conflicts. This mechanism can be conceived like a thermostat with the respective inputs and outputs connected by an effective feedback procedure in a cybernetic sense (see also Deutsch, 1963). Such relationships can be illustrated in a simplified system model (see Figure 3).

This system model should not, however, be equated with systems theory in a more demanding sense (e.g., Luhmann, 1984). Thus, such systems need not necessarily be in equilibrium and they may also explode or implode as, in fact, they did in Communist Eastern Europe.

Nevertheless, such a model is again helpful to locate the major subdivisions of politics (and political science), which also constitute the major subsections of this encyclopedia and, in fact, many political science departments or national associations. The bottom square includes, in a broader sense, the fields of political sociology and, when this is treated separately, political economy. The square on the left-hand side represents political sociology in a narrower sense of the term (organized interest groups, political parties, etc.). The top square reflects the institutional side (involving a possible separation of powers, etc.) but also questions of governance often including the realm of public policies and public administration on the right-hand side. All this is embedded in the international system concerning interactions with the outside

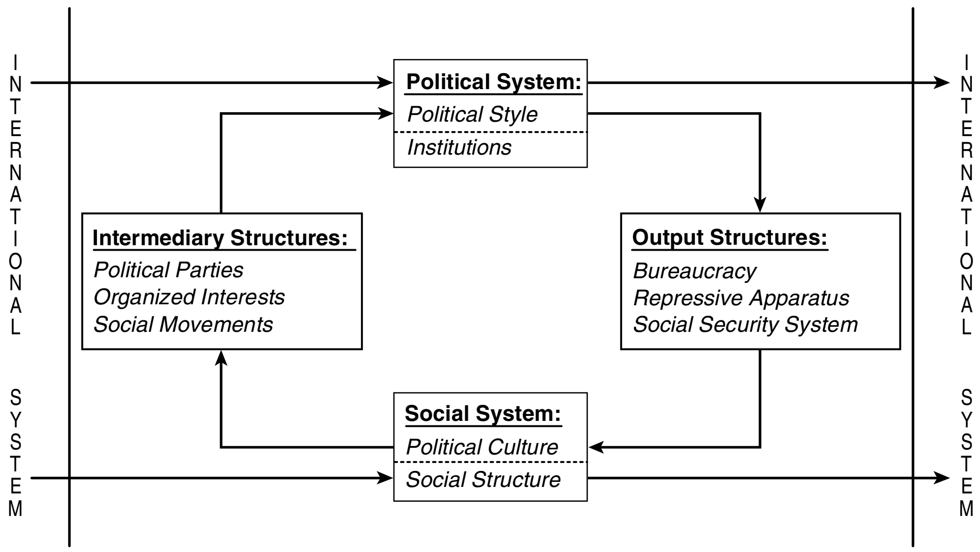


Figure 3 Simplified System Model

Source: Dirk Berg-Schlosser.

world both of state and society as the field of international politics and, in a more limited sense, international political economy. The arrows of such interactions can go in both directions. The systematic comparison of such systems or some subfields is the realm of comparative government. Overall theoretical (and philosophical) implications are the concern of political theory, and the respective methods and analytic techniques applied constitute the subfield of political methodology.

Linkages Between Levels of Analysis

A final building block to be considered here concerns the links between macro-aspects of entire political systems and their relationship with the micro-world of individual citizens and the meso-level of organisations in between. For this purpose, what has been dubbed "Coleman's bathtub" (Coleman, 1990, p. 8) is most helpful. Here, a given objective (structural) situation at the macro-level (on the upper-left-hand side in Figure 4) can be linked to the micro-level of individual subjective perceptions and values, which are then translated into concrete actions, possibly aggregated on the meso-level, and then leading to the outcome on the macro-level to be explained (upper-right-hand side). This relationship is illustrated in Figure 4.

It is important to note at this place that we do *not* imply for the individual actors, as is done for example in economics and rational choice theory, a specific logic of selection, as, for example, maximizing a person's material well-being. Such very restrictive assumptions of "homo oeconomicus" or even "homo sociologicus" (Dahrendorf, 1977) only rarely apply in political science, where usually a much wider range of choices exists, even if some of which may appear as "irrational" to others (for example, strongly felt ethnic or religious identities).

The purpose here, again, rather lies in the possibility to *locate* various approaches and their respective assumptions in such a scheme and to show the plurality of concepts that can be integrated here, but keeping them in a coherent relationship. Hartmut Esser (1993, p. 23 ff.), for example, has extended possible assumptions at the micro-level to include "restricted, resourceful, evaluating, expecting, maximing men" (RREEMM) or women, and even further assumptions about conflicting "identifying" (with some collective entities) or "individualizing" attitudes (RREEIIMM) or similar ones may be added. For the logic of the situation also framing procedures play a role where individual perceptions are shaped by the social milieus of one's childhood and later environment (see also D'Andrade, 1995). The point here is to

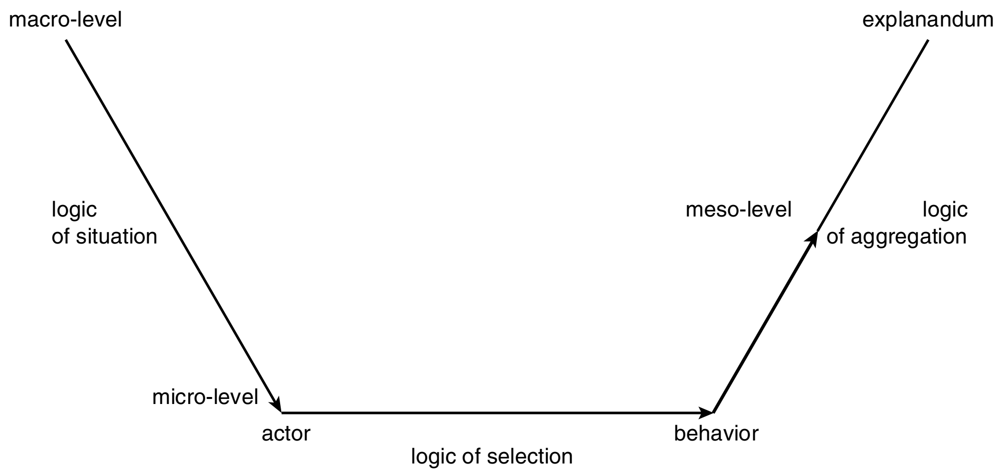


Figure 4 Linking Levels of Analysis

Sources: Adapted from Coleman (1990) and Esser (1993, p. 98).

show that in this way given historical and traditional constraints at the macro-level terms can be meaningfully linked to individual and, at the aggregated level, collective political action. Which particular historical, cultural, or other factors condition these choices in any given situation can be left open at this place, leaving room, again, for a plurality of cultural and theoretical perspectives.

The Need for a Reflective Pluralism

As this overview has shown, there are some basic building blocks, which can be usefully employed in a variety of ways for locating different epistemological positions and historical-cultural traditions in political science and similar fields. In this way, it at least becomes clearer where (and perhaps also why) certain contending positions actually differ. We do not intend to "harmonize" these positions. They all have, to varying degrees, their respective strengths and weaknesses, and no coherent, well-integrated theoretical building is constructed here with these blocks. That may even not be desirable, leaving some room to agree to disagree about some basic issues and perspectives. What is desirable instead is to elevate our consciousness and our way to deal with such controversies to a level of reflective pluralism, where not just anything goes, but where contending epistemologies and approaches can be brought into a fruitful interdisciplinary, intercultural, and, possibly even, meta-theoretical dialogue.

As already mentioned, political science has always been characterized by a diversity of contending meta-theoretical positions, paradigms, and approaches. In Europe, in the last century various strands of normative-ontological, Marxist, and empirical-analytical persuasions have been at the forefront (for such and similar terms see, e.g., Easton, Gunnell, & Graziano, 1991; Quermonne, 1996). For several decades in the United States, "behavioralist" positions and, more recently, "rational" and "public choice" approaches have dominated (cf. the influential volumes by King, Keohane, & Verba, 1994, and Brady & Collier, 2004). In other parts of the world, different theological, philosophical, and epistemological traditions have influenced the (more recent) emergence of political science there. Altogether, thus a great variety of contending positions, which have been summarized as "naturalist," "constructivist," and "realist," can be observed (Moses & Knutsen, 2007).

Relations With Other Social Sciences

Pluralism and different traditions in political science also emerge when we change perspective and focus more precisely on its relationships with other social sciences. This section explores three sources of political science as it differentiated itself from other disciplines.

Evolution of Political Science

When looking at the period after World War II, the basic difference in the traditions of different countries and areas of the world is between a plural form (political sciences) that is more common in Europe and encompasses the singular (political science). Conversely, in the tradition of United States the singular form (political science) includes the plural (political sciences). In the singular, there is a pluralist political science where empirical analysis is dominant, but also other perspectives (law, history, philosophy) are present. However, be it plural or singular, during the last decades empirical political science has increasingly differentiated itself from sociology, and above all from political sociology, public law, political philosophy, and contemporary history. Actually, in these developments we can see differences among disciplines or, more precisely, among specific groups of scholars in specific countries, but also overlapping and mutual influences with ever stronger interactions among scholars who are able to cross borders from Europe to North and South America, and to Africa and Asia, with a strong British tradition still present in Australia.

When we trace the original development of empirical political science, we can see that in a large number of European and American countries, political science is the result of empirical developments in public law. Consequently, the first difference concerns the difference between the perspective of law, which deals with "what ought to be"—with norms and the institutions that seek to embody them, and that of political science as transformed by behavioralism into an empirical social science, which is focused on "what is"—on the reality and on the explanations of it.

In Europe as well as in North and South America, there are other strong traditions that make contemporary history a parent of the new, post–World War II empirical political science. Here, despite all its ambiguities, the criterion of differentiation is between historical idiographic research, focused on the analysis of specific unique events, and a political science characterized by epistemological and methodological assumptions of other social sciences such as economics, sociology, and psychology, at least in terms of expectations of empirical findings (nomothetic) with a more general scope (regularities, patterns, laws). Social history and historical sociology as in the works of Reinhard Bendix, Barrington Moore, Stein Rokkan, Charles Tilly, and others have also greatly contributed to our understanding of long-term political processes at the macro-level. In this respect historical studies and political analysis can nicely supplement each other, as in the adage "Political science without history has no root, history without political science bears no fruit."

Within the European and North American traditions, sociology is the third parent of the new empirical science. Here, in addition to the common epistemology and possibly methodology of research, the overlapping of the contents, when political sociology is considered, makes the differentiation more difficult. Such a criterion was set up by two famous sociologists of the 1950s, Bendix and Seymour Lipset, when they stated that political science starts from the state and analyzes how it influences society, whereas political sociology starts from the society and analyzes how it influences the state (Bendix & Lipset, 1957, p. 87). In other words, the independent variables of a sociologist are the dependent variables of a political scientist: The arrows of explanation are going in opposite directions. Such a distinction sounds artificial and unrealistic when the inner logic of research is taken into account—if we decide in advance what is/are the independent variable/s, how can we stop when no salient results come out and declare that from now on one becomes a sociologist or economist or else? Nevertheless, for years such a distinction was the rule of thumb used to stress the difference between political sociology and political science. However, such a rule was responding more to the necessities of differentiation between academic communities than to the needs of developments in empirical research. It must also be noted that political sociology can be understood in both a broad and a narrow sense. In the former, it covers the broad social-structural and political-cultural bases of politics and their long-term developments over time at the macro level. In the latter, the intermediate and input structures of politics like interest groups, parties, social movements and other aspects of civil society are dealt with. These, undoubtedly, belong more to the realm of political science proper and have continued to flourish. In the former sense, closer to historical sociology, a certain slackening can be observed. This is due to the fact that the consideration of long-term social-structural developments had rigidified to some extent in the 1970s and 1980s in variants of orthodox Marxism, or the political element had largely disappeared in the analysis of finer social distinctions in Pierre Bourdieu's sense.

Last but not least, the development of differences between political philosophy and political science should be recalled. Again, there is much overlapping of contents, but epistemology and methods are different and easy to distinguish. As recalled by Giovanni Sartori (1984) with regard to the "language watershed," first of all, the language is different: The words and the related empirical concepts of political science are operationalized, that is, translated into indicators and, when possible, in measures, whereas the language of political philosophy is not necessarily so; it usually adopts meta-observed concepts, that is, concepts that are not empirically translated.

As discussed in the next section, this apparently simple differentiation covers possible commonalities, but leaves unsolved how the two different disciplinary perspectives deal with normative issues. Norberto Bobbio (1971, pp. 367, 370) made a relevant contribution in this direction when he emphasized that political philosophy focuses mainly on

- the search for the best government;
- the search for the foundations of the state or the justification of political obligations;
- the search for the 'nature' of politics or of 'politicness'; and
- the analysis of political language.

All four topics have an ethical, normative content, which is a characterizing feature of each political philosophical activity. At the same time, Bobbio recalls that an empirical analysis of political phenomena that are the objects of political science should satisfy three conditions:

- the principle of empirical control as the main criterion of validity;
- explanation as the main goal; and
- *Wertfreiheit,* or freedom from values, as the main virtue of a political scientist.

As noted in the discussion of epistemology above, the key element is in differentiating the speculative, ethically bound activity of a philosopher from the empirical analysis, even of phenomena that are influenced by the values of the actors.

The Influences of Other Disciplines

The obvious conclusion of the previous subsection is that there are different ways of analyzing political phenomena that correspond to different traditions and come from different cultural influences. Moreover, the discussion of those differences may help in developing a negative identity of political science. This is the very first meaning of the actual pluralism we have in this domain of knowledge: Pluralism only means that politics can be legitimately studied in different ways and with different goals that belong, at least, also to law, history, sociology, and economics. Pluralism in this sense challenges the autonomy of political science and even, in a radical version, has led to a denial that it constitutes a specific science. This view, however, no longer corresponds to the internal differentiation of the discipline, its specific achievements, and its more general institutionalization as an academic field. In addition, a second sort of pluralism inside political science proper reveals the overlapping and the influences of other disciplines in empirical political science. In this vein, when again considering the period starting after World War II, a main hypothesis can be proposed: Political science is influenced by the discipline or the other social science that in the immediately previous years has developed new salient knowledge. This is so for sociology, as can be seen in the analysis of Lipset and Bendix and other important authors since the end of World War II, who developed the work of classic sociologists, from Weber and Durkheim to Parsons and others. This is so for the influence of general systems theory, coming from cybernetics, and translated meaningfully into the analysis of political systems so that since the mid-1950s, it has become a major approach in political science. The same applies to the influence of functionalism, born with the developments of anthropology, and to rational choice or more specifically game theory, coming from economics and becoming more and more influential with several adaptations since the end of the 1950s. This is so, finally, for cognitive psychology that became very important in economics and at the same time in political science with the development of new ways of studying electoral behavior.

Moreover, when we consider more closely some of the subsectors of political science we can see more specific influences. For example, in the field of international relations, we can see the influence of international law. The public policies sector of political science has been influenced by sociology, economics, and constitutional and administrative law. Sociology has shaped the development of research on political communication. The influence of history can be seen in the selection of specific topics in comparative politics. Thus, we see that political science not only embodies a highly developed pluralism of the two kinds mentioned above, but also requires the integration of knowledge from other disciplines. Political scientists therefore also need an educational background that enables them to draw on these interdisciplinary sources of the field.

Recent Developments and Perspectives

As the International Political Science Association notes on its website,

> it is hard today for political scientists of 2011 to imagine the very different status their discipline in the world under reconstruction of 1949. In place of the familiar, well-structured web of national associations we know today, there were associations only in the United States (founded in 1903), Canada (1913), Finland (1935), India (1938), China (1932), and Japan (1948). (http://www.ipsa.org/history/prologue)

Founders of the International Political Science Association met in 1948 to plan for a new international organization that would establish dialogue among political scientists throughout the world. Taking into account the views of political science as variously defined in different countries, they identified four fields as constituting the discipline, acknowledging "the influence of the philosophers with 'political theory,' the jurists with 'government,' the internationalists with 'international relations,' and the fledgling behaviorist school of American political science with 'parties, groups and political opinion.'" Today IPSA serves as the primary international organization in the field, with individual and institutional members as well as affiliations with national political science memberships across the globe.

With respect to the ways that pluralism and interdisciplinary developments have taken place in political science, the North American influence has been paramount. The so-called Americanization affected all of Europe as well as other areas of the world where native scholars, educated in North American universities, went back to conduct research and to teach, bringing a new empirical conception of the discipline that significantly contributed to create new communities of political scientists (Favre, 1985). Moreover, American foundations and research centers gave support for research in Europe, Latin America, Asia, and Africa. While there are differences in political science as it exists today on different continents in this domain of knowledge—and actually also in most other scientific research domains—the North American universities, as well as the American research centers and the scholars associated with them, had a great influence that can be compared only to the intellectual German influence during the 50 years between the end of the 19th century and the first 3 decades of the 20th century. Thus, at the end of the 1960s, Mackenzie (1969, p. 59) suggested that in this period 90% of political scientists worked in North America, and Klaus von Beyme noted that the Department of Political Science at the University of California, Berkeley had more professors in this field than all German universities combined. Moreover, in those years and earlier in the 1950s in all European countries and in Japan, the American influence had been very strong in all social sciences, with some exceptions such as anthropology, which had a specific French presence. Forty years later, 70% of all political scientists are almost equally present in North America and Western Europe and the other

30% are spread throughout the rest of world, again with a relatively strong presence in Japan.

To better understand the development of the discipline all over the world with its specific contents, approaches, and methods, we should note that the American influence has been supplemented by the great increase of faculty members in all universities of the world since the 1960s. During this period, especially in Europe, there was the so-called transition from elite universities to mass universities; that is, there was a significant growth in the number of university students, which required the recruitment of a large number of new faculty members in all disciplines, political science included. This growth of the discipline allowed the creation of academic groups who absorbed and translated the American influence in different ways. Without that internal growth, there would not have been even the possibility of such a widespread influence.

This penetrating influence had a different impact in the various countries also in connection with their respective traditions. More precisely, on the one hand, the influence of the way empirical research is developed through quantitative statistical analysis and qualitative research is general and fairly homogeneously widespread; on the other hand, some approaches that have a stronger correspondence or congruence in the European and Japanese traditions, such as the different neo-institutionalist approaches, have had more success than other approaches, such as the rational choice approach. That latter has become very strong in North American political science, where it has its roots in economics, but it has remained much weaker among political scientists in other areas of the world. By its very nature, political science in other regions of the world also has been more specifically historical and comparative rather than just focusing (mostly) on a single case, the United States. Moreover, the legal traditions of several European countries especially influenced research in the subfield of public policies. At the same time, traditions in political philosophy and contemporary history maintained some influence on research that was predominantly qualitative rather than quantitative. Finally, and more specifically in Europe, research funding from the European Union led to the development of a number of works focused on topics related to the Union.

In the most recent developments, the impact of a more continuous and effective communication among scholars through different modalities, such as domestic and international collective associations, research networks, and initiatives of private and public institutions, affected the discipline as a whole mainly in three directions. The first one is a growing trend toward blurring national differences and a consequent convergence between North America, or between North and South America, and Europe. The second is an increased blurring of subdisciplinary divides. This is so especially between comparative politics and international relations, traditionally two separate fields in the past. Such a trend is particularly evident in the European studies. Third, research in political science more and more focuses on relevant, contemporary realities rather than confining itself to an ivory tower, which made it distant and largely irrelevant and, consequently, created that "tragedy of political science" Ricci singled out years ago (1987). Contemporary political science thus has developed into a multi-faceted, well-established discipline that is concerned with the pressing problems of our times and provides sound empirical analyses and meaningful orientation in the ever more integrated and complex world of the 21st century.

Bertrand Badie
Professor of International Relations,
Paris Institute of Political Studies (Sciences Po)
Paris, France

Dirk Berg-Schlosser
Emeritus Professor of Political Science,
University of Marburg
Marburg, Germany

Leonardo Morlino
Professor of Political Science, LUISS
Rome, Italy

See also African Political Thought, Behavioralism; Buddhism, Comparative Politics; Concept Formation; Constructivism; Epistemological and Methodological Foundations; Greek Philosophy; Hinduism; History of Political Science; International Relations as a Field of Study; Islam; Marxism; Policy Analysis; Political Sociology as a Field of Study; Political Theory

Further Readings

Adorno, T., Albert, H. Dahrendorf, R., Habermas, J., Pilot, H., & Popper, K. R. (Eds.). (1969). *Der Positivismusstreit in der deutschen Soziologie* [The positivism debate in German sociology]. Darmstadt, Germany: Luchterhand.

Almond, G. A., & Genco, S. (1977). Clouds, clocks, and the study of politics. *World Politics*, *29*(4), 489–522.

Aristotle. (1962). *Politics*. Harmondsworth, UK: Penguin. (Original work 350 BC)

Bendix, R., & Lipset, S. M. (1957). Political sociology: An essay and bibliography. *Current Sociology,* 2(VI):79–99.

Bobbio, N. (1971). Considerazioni sulla filosofia politica [Considerations on Political Philosophy]. *Rivista Italiana di Scienza Politica, 1*(2), 367–380.

Brady, H. E., & Collier, D. (2004). *Rethinking social inquiry: Diverse tools, shared standards.* Lanham, MD: Rowman & Littlefield.

Clastres, P. (1975). *La société contre l'Etat.* [Society against the state]. Paris: Minuit.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.

D'Andrade, R. G. (1995). *The development of cognitive anthropology*. Cambridge, UK: Cambridge University Press.

Dahrendorf, R. (1977). *Homo Sociologicus*. Opladen, Germany: Westdeutscher Verlag.

Deutsch, K. W. (1963). *The nerves of government*. New York: Free Press.

Durkheim, É. (1975). *Textes*. Paris: Editions de Minuit.

Easton, D. (1965). *A systems analysis of political life.* New York: Wiley.

Easton, D., Gunnell, J. G., & Graziano, L. (Eds.). (1991). *The development of political science*. London: Routledge.

Esser, H. (1993). *Soziologie. Allgemeine Grundlagen* [Sociology. General Foundations]. Frankfurt am Main, Germany: Campus.

Evans-Pritchard, E. (1940). *The Nuer*. Oxford, UK: Clarendon Press.

Favre, P. (1985). *Histoire de la science politique* [History of political science]. In M. Grawitz & J. Leca (Eds.), *Traité de science politique* [Treatise of political science] (Vol. 1, pp. 28–41). Paris: Presses Universitaires de France.

Foucault, M. (1970). *The order of things*. London: Routledge.

Geertz, C. (1973). *The interpretation of cultures.* New York: Basic Books.

Goodin, R. E., & Tilly, C. (Eds.). (2006). *Contextual political analysis*. Oxford, UK: Oxford University Press.

Habermas, J. (1975). *Legitimation crisis.* Boston: Beacon Press.

Habermas, J. (1981). *Theorie des kommunikativen Handelns* [Theory of Communicative Action] (2 vols.). Frankfurt am Main, Germany: Suhrkamp.

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science.* New York: Free Press.

Jackson, R. (2000). *The global covenant*. Oxford, UK: Oxford University Press.

Kant, I. (1956). *Kritik der reinen Vernuft* [Critique of pure reason]. Hamburg, Germany: Felix Meiner. (Original work published 1787)

King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry*. Princeton, NJ: Princeton University Press.

Lasswell, H., & Kaplan, A. (1950). *Power and societies*. New Haven, CT: Yale University Press.

Lorenzen, P. (1978). *Theorie der technischen und politischen Vernunft* [Theory of technical and political reason]. Stuttgart, Germany: Reclam.

Luhmann, N. (1984). *Soziale Systeme* [Social systems]. Frankfurt am Main, Germany: Suhrkamp.

Lukács, G. (1967). *Geschichte und Klassenbewußtsein* [History and class consciousness]. Darmstadt, Germany: Luchterhand.

Mackenzie, W. M. (1969). *Politics and social science*. London: Penguin.

Madan, T. N. (Ed.). (1992). *Religion in India*. Delhi, India: Oxford University Press.

Mbiti, J. S. (1969). *African religions and philosophy*. London: Heinemann.

Merton, R. K. (1968). *Social theory and social structures*. New York: Free Press.

Monroe, K. R. (Ed.). (2005). *Perestroika! The raucous rebellion in political science.* New Haven, CT: Yale University Press.

Morgenthau, H. (1948). *Politics among nations*. New York: Knopf.

Moses, J. W., & Knutsen, T. L. (2007). *Ways of knowing: Competing methodologies and methods in social and political research*. London: Palgrave Macmillan.

Parsons, T. (1951). *The social system*. Cambridge, MA: Harvard University Press.

Plato. (1945). *The republic*. New York: Oxford University Press. (Original work 360 BC)

Popper, K. (1972). *Objektive Erkenntnis. Ein evolutionärer Entwurf* [Objective knowledge: An evolutionary approach]. Hamburg, Germany: Hoffmann und Campe.

Quermonne, J.-L. (Ed.). (1996). *Political science in Europe: Education, co-operation, prospects, report on the state of the discipline in Europe*. Paris: Institut d'Etudes Politiques.

Ricci, D. (1987). *The tragedy of political science: Politics, scholarship, and democracy*. New Haven, CT: Yale University Press.

Sartori, G. (Ed.). (1984). *Social science concepts: A systematic analysis*. Beverly Hills, CA: Sage.

Shin, D. C. (1999). *Mass politics and culture in democratizing Korea*. Cambridge, UK: Cambridge University Press.

Weber, M. (1968). *Economy and society*. New York: Bedminster.

A

ACCOUNTABILITY

Accountability has become a key concept in both public administration and democratic theory. Its meaning is contested, but the general definition "obligation to answer for the performance of duties" would fit most versions. In this sense, accountability is a relationship between two parties—the person or organization answering or being held to account (the accountor or agent) and the person or organization to whom the account is owed (the account holder or principal). Analysis of accountability therefore begins with the double question: *Who* is accountable *to whom*? Accountability obligations depend on the terms of the relationship and on its institutional context, leading to two more key questions: (1) *For what* is the accountor accountable and (2) *how*? This entry covers the following topics: defining accountability, typologies of accountability, mechanisms of public sector accountability, democratic accountability, accountability in international relations, single versus multiple accountability, accountability in networks, and accountability and the new public management (NPM).

Though the English word *accountability* has a respectable historical pedigree (the *Oxford English Dictionary* records its use from the late 18th century), its prominence in political science dates only from the 1980s, before which time the cognate term *responsibility* was preferred. (Indeed, linguistic equivalents of "responsibility" are still dominant in other European languages that lack a direct parallel for "accountability.") The rapid rise of "accountability" can be traced to its adoption by public choice theory, particularly principal–agent theory, and by management theory, which in turn greatly influenced democratic theory and public administration. However, the relative lack of intellectual history and of cross-disciplinary seminal texts has meant that academic analysis of accountability has proceeded in a haphazard, fragmented, and repetitive fashion. Different subdisciplines, including comparative politics, international relations, deliberative democracy, constitutional law, and public management (further subdivided into U.S. and European versions) have each been developing their own parallel theories of accountability, with little cross-fertilization or sense of common purpose.

Defining Accountability

Though the core sense of accountability, the obligation of the accountor to answer for the performance of duties, is uncontroversial, disagreement occurs over what should be added to that core. Most analyses also include the capacity of the account holder to impose sanctions or other remedies on the accountor as a necessary complement to full accountability. However, some versions confine accountability to the initial informing and discussing stages, omitting the requirement for any rectification. In effect, they equate accountability with *transparency*, another popular term with which accountability is frequently linked. Certainly, in complex modern systems of public accountability, some accountability mechanisms, such as parliamentary inquiry or media investigation, can provide transparency but lack the capacity for imposing sanctions, leaving that function to other agencies such as courts or the executive. But transparency on its own, with no prospect of correction or other adverse consequences, falls short of full accountability.

In its core sense, accountability, like accounting itself, is essentially retrospective or *ex post*, in that it is concerned with information and explanation about past actions of the accountor. Discussion about future actions is, therefore, not, strictly speaking, an exercise in accountability. However, in a continuing relationship between principal and agent, *ex post* can easily overlap with *ex ante*, as in election campaigns when incumbent representatives not only defend their past actions but also outline their future plans. Indeed, some theorists have wanted to distinguish two types or aspects of accountability, *ex post* (retrospective) and *ex ante* (prospective) accountability, while others have included all communication between political leaders and the public as part of an ongoing accountability dialogue. Such usage, while understandable because of the obligations of democratic governments to engage in continuing discussion with their citizens, extends accountability beyond its normal focus of answering for past actions. Similarly, government consultation with stakeholders about future policy certainly helps keep politicians and bureaucrats in touch with relevant sections of public opinion, but it does not necessarily imply accountability in the strict sense of accounting for previous decisions.

Accountability is also sometimes taken beyond its retrospective core when it is equated with institutional devices for limiting or constraining power. For instance, constitutional checks and balances, such as federalism and the separation of powers, are sometimes described as mechanisms of accountability because they limit the legal power of governments and prevent them from abusing the rights of citizen. Such constraints may certainly involve accountability mechanisms, for instance, when a government oversteps its legal powers and is called to account by the courts. However, constitutional laws and regulations themselves are essentially prospective in focus, seeking to limit and control future actions of governments. The mere fact that laws and regulation constrain the power of governments need not in itself imply that they are instruments of accountability.

Significant disagreement also surrounds how far accountability is to be distinguished from other, closely related concepts, namely, responsibility and responsiveness. "Responsibility" and "accountability" share parallel conceptual histories, both originating in the notion of answering ("responding," "giving an account to") to someone. "Accountability," as the obligation to answer for the performance of duties, generally implies a relationship between two or more parties, in which one party is subject to external scrutiny from others. "Responsibility," on the other hand, typically refers to the internal deliberations and actions of a single person or organization.

This division of conceptual labor is by no means universally observed. "Responsibility" sometimes includes answering to others as well as making individual choices, as, for instance, in the classic British conventions of "responsible government" and "ministerial responsibility." Indeed, before the rise of the term *accountability*, *responsibility* regularly covered both external and internal aspects (as its equivalents still do in other European languages).

The popularity of "accountability" has seen it extended in a similar fashion, but in the opposite direction, to include not only external scrutiny but also the internal capacity for considered and conscientious action usually described as acting "responsibly." For instance, where members of caring professions, such as social workers and health professionals, have a strong vocational commitment to serving the interests of the community, this concern is sometimes identified as in itself constituting accountability to the community, regardless of whether members of the community actually have any rights of scrutiny or complaint. Similarly, members of the nonprofit charitable sector commonly see themselves as accountable to the recipients of charity even though they are not, strictly speaking, answerable to them. Particularly where professional standards or values are publicly articulated, voluntary adherence to such standards can be seen as a form of accountability without any requirement to report externally or without external any mechanism for others to demand compliance. The accountability has become internalized, as though the conscientious professional is in an imaginary dialogue with his or her clients and answering to them.

The question of whether an internalized sense of service should be accepted as a form of accountability recalls the perennial debate over the relative merits of professional discretion compared with external scrutiny as a means of securing public service in the public interest. The issue was classically discussed by Carl Friedrich and Herman Finer in the 1940s as a debate over two contrasting types of responsibility but is now typically described as a clash between two types of accountability. Conceptual clarity might be better served if accountability were reserved for external scrutiny and rectification, recasting the Friedrich–Finer debate as being between the respective merits of (internal) responsibility and (external) accountability. However, the emotive pull of "accountability," like that of "democracy," makes supporters of professional discretion unwilling to accept what would amount to an accountability deficit. Instead, they prefer to see themselves as embracing a special form of accountability.

Similar extensions beyond the core meaning of accountability derive from the connection between accountability and responsiveness. Accountability, as the obligation to answer to external scrutiny, derives its main justification from its contribution to responsiveness, understood as the readiness of institutions and officials to respond to the needs and interests of those whom they serve. In public administration, for instance, responsiveness, the alignment between official action and public preferences, is the goal to which the accountability of governments is a key means. However, accountability procedures of external scrutiny, such as review and audit, are not the only mechanisms for making governments responsive. For instance, changing the organizational culture of a public agency toward a greater client focus and concern for service quality can make the agency more responsive to the public but need not involve additional accountability procedures of scrutiny or inquiry. Yet many in such an agency would typically claim that it had become both more responsive and more accountable, in effect identifying accountability with responsiveness.

The link between accountability and responsiveness also arises with market mechanisms. Markets are unquestionably instruments of responsiveness. They find their main rationale in their capacity to align the provision of goods and services with the preferences of consumers and citizens. But are they therefore instruments of accountability? According to a strict understanding, organizations operating in a market, particularly private companies, are accountable primarily to their owners and shareholders, not to their customers. They may be accountable to individual consumers who have purchased goods or services, and they may also be publicly accountable in the sense of being liable to scrutiny for complying with any relevant law and regulations. But overall responsiveness to consumer preferences comes from the consumers' capacity to choose between alternative suppliers in a competitive market. Suppliers adjust to consumer demand not because of any complaints or scrutiny from potential customers, who have no such rights against a supplier, but because they will go out of business if they have no customers.

Markets are thus primarily "exit" mechanisms for securing responsiveness. Dissatisfied customers vote with their feet. Accountability, on the other hand, as normally understood, is a "voice" mechanism, allowing dissatisfied members of the public to complain and to seek information and rectification from an organization. In this case, markets do not count as accountability mechanisms. Nonetheless, because competitive markets have the capacity to force service providers to take note of consumer preferences and, in some sense, to "answer" to expressed demand, to describe them as instruments of accountability can have a certain plausibility.

In the same way, organizations that respond to peer or public opinion through concern for their corporate reputations are sometimes said to be exhibiting "reputational" accountability. Again, there is no direct connection or dialogue between the supposed "accountor" and the "account holder," no right of the public or the organization's peers to call the organization to account, and no formal obligation of the accountor to accept sanctions or redirections from others. Again, however, the emotive force of the term *accountability* leads sympathetic observers to classify such responsiveness as instances of accountability. Moreover, in modern democratic societies, reputational effects are often the result of media publicity and scrutiny. There are good grounds for seeing the media as agents of accountability, holding public figures and organizations up to scrutiny, even though the media may have no formal rights to demand information or to impose sanctions. As with transparency in general, however, such accountability without rectification is inchoate and incomplete.

Accountability is thus a chameleon-like concept that readily takes on new, additional senses from the different contexts in which it is used. Because external scrutiny and sanctions, the core of accountability, are so central to checking abuses of power and to the processes of representative democracy, the term itself is easily extended to other mechanisms and processes that secure the same overall objectives, including legal and regulatory constraints, market competition, and public service professionalism and commitment to the public interest. These extensions and variations have been driven by the emotive power of the term combined with the lack of widely recognized academic authorities on the topic. To expect agreement on a single concept of accountability is unrealistic. But, at least, analysts of accountability could become more aware of the well-established variations in usage and more willing to place their own versions within that larger conceptual context.

Typologies of Accountability

Most analyses of accountability divide it into different types, though, as with the definition of accountability, there is no agreement on a typology or even on the meaning of some of the labels. Some typologies are based on distinctions in the subject matter of accountability (*for what*), for instance, between accountability for contestable outcomes ("political") or for agreed tasks ("managerial") and accountability for different types of activity, such as finances, processes, and performance. Other typologies focus more on the institutions and mechanisms of accountability, for instance, "political," "legal," "bureaucratic," and "professional."

Several classifications are built on distinctions in the direction of accountability, such as "vertical," which can include "upward" accountability within a structured hierarchy as well as downward to citizens and customers, and "horizontal" and "outward," which refers to accountability to institutions or individual of roughly equal status. A "360-degree" accountability implies accountability in all directions: upward, outward, and horizontal. "Internal" versus "external" accountability can refer to the difference in *to whom* an organization is accountable, internal referring to those who exercise clear ownership or delegation rights while "external" refers to those who have no such rights but are nevertheless affected.

Typologies tend to be context dependent. For instance, writers on developmental politics and democratization have employed the contrast between "horizontal" and "vertical" accountability to stress the importance of having executive governments accountable (horizontally) to other coequal institutions, such as courts, legislatures, and auditors. Horizontal accountability has become equated with the rule of law and constitutional government, seen as prerequisites for successful representative democracy. In international politics, on the other hand, types of accountability identified in one influential study (hierarchical, supervisory, fiscal, legal, market, peer, and reputational) reflect the political realities of international relations and the absence of some of the more robust accountability mechanisms, such as elections and legal sanctions, available within nation-states.

In the public administration literature, the structure of accountability typologies tends to follow constitutional structure, with a sharp division between the approach in the United States and that in the parliamentary systems of the United Kingdom (UK); other Westminster systems such as Canada, Australia, and New Zealand; and Western Europe. In the United States, the separation of powers, along with federalism, makes bureaucrats independently accountable to a range of institutions, including the president, Congress, and the courts, forcing them to juggle between competing accountability demands. Analyses of accountability tend to stress the discretionary role of bureaucrats and the varied types of accountability forum—hierarchical, political, legal, and professional—in which they operate.

In parliamentary systems, by contrast, bureaucratic accountability centers on the hierarchical chain of accountability through ministers to parliament and the public, associated with the traditional conventions of ministerial responsibility and described variously as "parliamentary" or "political." To this have been added other, supplementary types, such as "legal," "judicial" (to cover accountability through courts and tribunals), or "managerial," referring specifically to the output-focused accountability emphasized in the NPM.

Not only do different typologies contain different sets and combinations of individual types of accountability but particular individual types also vary in meaning between different typologies. Political accountability, for instance, is sometimes confined to politicians and to processes involving politicians while at other times it refers, more broadly, to contestable outcomes involving non-elected officials as well as elected politicians. Legal and judicial accountability typically imply the processes of courts and the legal system. But they vary as to whether they are confined to the judicial branch of government or also include quasi-judicial tribunals or other executive-based institutions such as ombudsmen and freedom of information rules. Legal accountability has also been defined even more broadly to cover external monitoring for compliance with legally established rules, which therefore implies that legislatures scrutinizing compliance with their own legislation are engaging in legal accountability.

Finally, professional accountability also has no settled meaning. Sometimes it refers to the accountability that members of an expert profession owe to each other (peer accountability). For example, some professionals, such as doctors and lawyers, can be held to account by disciplinary bodies drawn from the profession itself. Less formally, members of a profession, including public servants, may be monitored and called to account through shared networks and collegial relationships. However, professional accountability has also been applied to the independent exercise of ethical norms of a particular profession, which have been internalized through socialization into the profession. In this case, professional accountability is being identified as a sense of professional responsibility, quite detached from any external scrutiny.

Mechanisms of Public Sector Accountability

The complex accountability structures surrounding modern governments involve a wide range of institutional mechanisms, which vary according to the section of government on which they focus (*who*?), the type of issue that they investigate (*for what*?), and the procedures they follow (*how*?).

Elections

In a representative democracy, the basic accountability mechanism is the general election at which incumbent executive leaders present themselves to the voters and seek a renewal of their mandate to govern. Elections compel elected politicians to explain and justify their actions and give the citizens the opportunity to listen and impose a verdict. Elections legitimize the control exercised by political leaders over the executive bureaucracy and thus underpin the accountability of agency heads to ministers and of junior officials to their superiors through the hierarchical chain of command. Though elections may be held at infrequent intervals, their indirect influence on government accountability is immense, through the threat of future of retribution imposed on unpopular governments.

Elections may be powerful but they are also blunt. Their effectiveness as accountability instruments can be compromised by their forward-looking function of selecting an incoming government. Voters may be deflected from sanctioning an unpopular government by the perception that the alternative leadership is even less palatable. Accountability is also denied where incumbent leaders do not stand for reelection (most notably in the case of second-term United States presidents). Moreover, because elections require a general judgment over a whole regime and its overall program, they do not allow for more fine-grained accountability on particular issues or decisions. Electioneering is dominated by general slogans and misleading rhetoric and does not offer much scope for accurate information or serious discussion.

Elections therefore need to be supplemented by a range of other accountability mechanisms and should not be seen as the sole instruments of democratic accountability, as is the tendency in theories of democracy built around the single act of voting and electoral choice. Conversely, regimes that lack elections, though undemocratic, may still offer citizens opportunities to hold their governments to account, for instance, through legal processes or complaints procedures.

Legislative Scrutiny

Between elections, the major institution of accountability is the legislature. Though primarily defined in terms of their lawmaking functions, modern legislatures have ceded much of their legislative initiative and discretion to the executive branch, especially in parliamentary democracies where the main legislative chamber is typically under the control of the ruling party. In compensation, legislatures have increasingly emphasized their accountability role as the main forum where the executive is required to answer to the public.

Legislatures hold executives to account through a variety of avenues. One is the requirement for regular reporting on executive activities. All government agencies and statutory officials are obliged to report periodically (usually annually) to the legislature on their general performance. Information provided to the legislature thereby enters the public realm and is available for general debate and discussion.

Legislatures also have the right to question members of the executive and subject them to public scrutiny. In Westminster-based democracies, conventions of ministerial responsibility require ministers to answer to parliament for the conduct of their departments, to provide information about decisions, and, where necessary, to impose remedies. Similar conventions exist in most parliamentary democracies where ministers are accountable to parliament. Ministerial responsibility has been a topic of perennial controversy, mainly because of a widespread but erroneous belief that it requires ministers to take personal responsibility for the actions of subordinates and resign when major mistakes are discovered to have been made within their departments. Such "sacrificial" resignation is rarely, if ever, undertaken and the doctrine of "vicarious" responsibility on which it is based (that ministers are personally responsible for the acts of subordinates) does not accord with practice. When ministers resign, they do so because of personal faults of their own making, such as misleading parliament or engaging in improper or corrupt conduct, and the decision is usually politically determined in terms of minimizing damage to the government's popularity.

Ministerial resignation is a side issue in assessing the role of ministers' responsibility in government accountability. Much more important, and the basis of the effectiveness of ministerial responsibility as an accountability mechanism, are the obligations it imposes on ministers to answer to the public on matters of public concern, either in parliament or directly through the media. Though ministers are not required to take personal responsibility for all actions of their subordinates, they must provide information and justification when asked. Refusal to respond publicly, which is common among leaders of private organizations, is politically unacceptable for elected politicians. At the same time, however, while avoiding outright deceit (which remains a strong ground for forced resignation), ministers can readily prevaricate and avoid disclosure of embarrassing information. Moreover, their monopoly of the right of public response can carry the corollary that departmental officials remain out of the public eye, which can shield them from legitimate public scrutiny.

To circumvent the problem of bureaucratic anonymity, legislatures also question public officials directly, usually through a system of legislative committees. The practice of committee scrutiny is most highly developed in the U.S. Congress, which, through the separation of powers, is an active partner in shaping government policy and has a legitimate interest in overseeing the whole conduct of government business. In parliamentary systems, the scope of questioning is somewhat more confined, exempting appointed officials from answering on issues of government policy, out of deference to the democratic mandate of the elected leaders. Even so, restricting questions to administrative matters for which officials are more immediately responsible still allows considerable opportunity for holding the bureaucracy to account.

Courts

All governments are subject to legal accountability through the courts because courts determine whether the government has acted within the law. The operation and effect of this power vary with a country's legal and constitutional structure. One contrast is between Anglo-American countries, where cases involving the government are heard in the same courts as civil cases, and some European countries, notably France, where a completely separate court structure is reserved for cases involving the state. Another contrast concerns the scope of judicial review. Where the constitution, as in the United States, defines and limits the powers of both Congress and the president, the courts become a forum for holding the government generally accountable across a wide range of substantial policy issues. On the other hand, in parliamentary democracies such as the UK, where few constitutional limits are placed on legislative power, opportunities for challenging policies through the courts are much more limited.

Most legal cases involving the government are brought by individual citizens and deal with particular decisions affecting them that have been made by government agencies. It is open to the court to rule whether a decision was taken within the powers legally conferred on the government agency; whether the citizen received natural justice, in terms of fair procedure and due process; and whether the decision itself was reasonable. Whether courts can decide on the actual substance and merits of a decision varies according to the provisions of the individual legal system. Some systems make use of quasi-judicial tribunals, which operate like courts though with a generally more relaxed approach to procedure and which, being technically part of the executive branch, are empowered to amend executive decisions.

Legal accountability, like litigation generally, suffers from being slow and expensive and is beyond the reach of most people for most issues. However, its availability as a last resort is crucial to the public's capacity to hold governments to account. Like the rule of law itself, of which it is a key element, accountability through an independent and honest judiciary is the foundation of all public accountability.

Auditors and Other Monitoring Agencies

Governments are overseen and investigated by a range of special-purpose accountability agencies. Of these agencies, the most long-standing are the offices of government auditors (variously described as "auditors general" and/or "comptrollers general"). Their traditional function has been the monitoring of government finances on behalf of the legislature to see whether public revenue and expenditure have been managed according to legislative authorization and according to standards of public probity and propriety. The historic function of "regularity auditing" for financial compliance has more recently been supplemented by "performance" ("value for money," "efficiency," "comprehensive") auditing that extends to assessing the efficiency and effectiveness of government programs. Performance audits usually take a program's objectives as given and then examine whether these objectives have been achieved and at what cost. In performance auditing, as distinct from regularity auditing, auditors typically lack any powers of sanction or rectification and can simply recommend changes and improvement.

In general, government auditors have proved essential in maintaining financial integrity in governments. Conversely, the absence of effective audit is a key indicator of weak and corrupt government systems. They have also been very successful in exposing bureaucratic waste and inefficiency. Even though they cannot mandate any remedies, the authority of their recommendations and the adverse publicity attached to the exposure of serious waste and inefficiency are often sufficient to prompt governments to follow their recommendations voluntarily.

Besides auditors, other investigating bodies include government inspectors and ombudsmen. Inspectors are officials established within particular government departments and agencies with the function of improving efficiency and effectiveness. Inspectors have been employed to monitor school and prison systems as well as government departments in areas such as taxation, defense, and security where bureaucratic performance is a matter of particular concern. The position of ombudsman, first introduced in Scandinavia, has been adopted worldwide as an avenue of complaint for individual citizens seeking redress in connection with particular decisions. Ombudsmen usually have the power to investigate and recommend but not to impose remedies. In spite of this limitation, however, they have proved an effective accountability mechanism midway between individual complaint and full legal proceedings.

Recent decades have witnessed an "audit explosion" as governments and government agencies become subject to increasing supervision by regulatory agencies. With the transfer of responsibility for providing public goods and services away from government departments under political direction to various forms of arm's-length providers, regulation has tended to replace political and bureaucratic direction as the means of making public service providers publicly accountable. A plethora of regulatory agencies now monitor different areas of public service provision, such as health and education, or different aspects of government activity, such as occupational health and safety or human rights. While most regulatory bodies are public bodies, some are privately established but have been granted legal powers—for instance, some consumers' associations and animal protection societies. Other private monitoring bodies have no legal mandate but operate more informally as observers and critics of government activities. These include a number of private international watchdogs, such as Greenpeace and Amnesty International and the financial ratings agencies, Standard and Poor's and Moody's, which have a great impact on economic policy because of the impact of their ratings decisions.

Freedom of Information and the Media

The public availability of information held by governments may be only the initial stage of a full accountability process and needs to be followed up by discussion and then, if necessary, by rectification. However, once such information is released into the public realm it can readily be used to spark a political reaction and force governments into remedies. Public access to government information is therefore an essential component of government accountability and is provided by a number of channels.

One such channel is the right of the general citizen to seek access to information, both personal information held about them as individual citizens and general information about government policy. Rights of freedom of information are found in most established democracies, with the United States having led the way and the UK being a reluctant latecomer. Certain exemptions usually apply, on grounds such as national security, cabinet confidentiality, commercial confidentiality, and protection of legal proceedings. Financial charges can also be imposed, especially on matters of general interest, with the result that information is generally sought by well-resourced journalists or by organized groups with political interests rather than by individual citizens.

The various media outlets, both print and electronic, also help spread information and stimulate debate. In part, their function is strictly intermediary, relaying to a wider public news items and arguments supplied by others. However, they also play an independent role in instigating inquiries and conducting investigations. Though, for the most part, privately owned and not formally part of the machinery of government, the media are essential to effective accountability in large-scale modern states. Indeed, a free press, along with elections and an independent judiciary, has been acknowledged as one of the key institutions in securing an accountable government.

Intra-Organizational Accountability

In addition to being externally accountable, government agencies, like all organizations, also exhibit internal structures of accountability whereby different members or sections are accountable to others within the organization. Indeed, from the perspective of individual officials, organizational accountability upward through the chain of bureaucratic command is often the most immediate and salient form of accountability in their daily activities. To assist in reinforcing upward accountability, organizations often impose their own in-house versions of independent scrutiny, such as internal audit or inspection, which mirrors and anticipates the financial monitoring of external auditors and inspectors.

Accountability of individual officials to their superiors is an essential element in the overall democratic accountability of executive government. If elected leaders are to be accountable for the actions of bureaucrats, then they must be able to rely on their bureaucrats' willingness to take direction and answer for their actions. Accountability upward is thus a corollary of control downward. Conversely, any slippage in control from above is typically reflected in lack of accountability upward.

Weakness in upward accountability may be symptomatic of the well-known propensity of bureaucrats to pursue their own policy agendas and apply their own judgment against the wishes of their superiors. It may also come from contrary accountability pressures. "Bottom-up" views of administration legitimate the direct accountability of street-level bureaucrats to their clients among the public and undermine the authority of organizational superiors. From this perspective, accountability through the "top-down" chain of bureaucratic command becomes an unwelcome constraint to be managed rather than a mandate to be respected. Bottom-up versions of accountability can claim their own democratic legitimacy through answerability directly to the members of the public rather than indirectly through the public's elected representatives.

Systems of Democratic Accountability

Democratic accountability can be viewed from various perspectives, from that of the citizen wishing to hold his or her government to account, from that of the elected politicians responding to the voters, or from that of the government official answering to political masters and to the public. Depending on one's position and interests, different accountability mechanisms claim the most attention. Politicians are focused primarily on elections and the media, whereas public servants are more concerned with accountability to superiors in the hierarchy. Citizens in strife with government agencies will be looking to grievance procedures and ombudsmen. In all cases, the structure of available accountability mechanisms appears complex and untidy, incapable of being reduced to tidy diagrams or flowcharts without oversimplification. Structures of accountability are inherently pluralistic and are better described as "webs," or possibly "systems," to indicate their complexity and fluidity.

The various accountability agencies and processes in webs of democratic accountability exhibit widely different functions. Some accountability processes, such as elections, focus on overall performance of government, while others, such as financial audit, concentrate on details of administration. Politicians and the media home in on politically controversial and sensational issues. Auditors and inspectors, by contrast, tend to deliberately steer clear of political controversy and instead concentrate on more humdrum areas of administration, which may appear dull but are often the site of major inefficiencies.

Some agencies, such as courts, exercise the full range of accountability functions, including information, discussion, and rectification, while others, such as the media, are limited to the transparency functions of information and discussion. It is for this reason that democratic accountability cannot be accurately represented by a chain of principal–agent relationships, given such relationships always imply the principal's right to impose sanctions. Standard structures of democratic accountability certainly involve account holders with the power to impose remedies and sanctions. This power may be located with a number of different actors. The basic right of rectification lies with citizens as voters who have the capacity to remove their elected representatives. But other public officers and institutions can also compel compliance on those accountable to them. Elected political leaders can direct their departments and agency heads in relation to their subordinates. Courts and tribunals can issue binding decisions.

However, not all agencies of accountability possess the power of rectification. Indeed, many of the most effective accountability agencies, including legislative committees, auditors, ombudsmen, and the media, can only investigate and recommend. Their effectiveness depends on their capacity to exert pressure on account holders with rectificatory powers, particularly executive leaders, to impose remedies in response to adverse publicity, backed by the ultimate electoral sanction. Democratic accountability can thus operate as an interlocking system with some institutions specializing in information and investigation while others are relied on for rectification. While there may be grounds for identifying elections as the defining accountability mechanism in representative democracies, they are far from being the only such mechanism.

A key factor in any web of accountability is whether each accountability agency charged with holding the government to account is itself publicly accountable for performing its accountability functions. Legislators, judges, auditors, ombudsmen, regulators, and so on need to be subjected to scrutiny themselves to prevent them from becoming lax or corrupt. Most officials in such scrutinizing positions are granted a degree of statutory autonomy to help them maintain an independent stance against the executive power that they must call to account. At the same time, however, this independence is itself open to abuse. Transparency, media scrutiny, and the ultimate power of arraignment and dismissal are essential buttresses of professional integrity. The guardians must be guarded.

Accountability in International Relations

Most analysis of government accountability has taken place within the context of nation-states, where the defining element of legal and political sovereignty provides a framework of effective sanctions to support other accountability mechanisms and where elections provide the people with the ultimate sanction. In the international sphere, however, the absence of effective sovereign power and elections raises the specter of an accountability vacuum. How can international organizations such as the United Nations (UN) or the World Bank be held accountable for serving the interests of the international community if that community is not capable of enforcing sanctions on them?

In practice, accountability in the international sphere, though weaker than in many nation-states, is far from nonexistent. In some areas such as trade policy and criminal behavior, international law can offer effective remedies backed up by the coercive power of nation-states. International organizations are subject to accountability mechanisms, such as audit, review, and media scrutiny, as well as being answerable to the national governments that underwrite them. The UN bureaucracy, though often incompetent and corrupt, can be exposed to detailed investigation, as over its handling of the Oil-for-Food Programme, and can be pressured to institute reforms. Rectification tends to be ineffective, because of the apathy and disunity among leading members of the Security Council, the body to which the bureaucracy is accountable. But, with more determination from member states, the accountability deficit could be significantly reduced. As it stands, the accountability deficit in the UN bureaucracy is probably no greater than that found in the national bureaucracies of most of the member states themselves.

The inherent contrast between accountability in nation-states and in the international arena should therefore not be overstated. Both are pluralistic in structure and involve a range of accountability mechanisms, many of which, such as the media and NGO (nongovernmental organization) watchdogs, operate at both levels. The significant lack of effective legal and political sanctions at the international level must be conceded. But so too must the potential for improved international accountability short of a united system of world government.

Single Versus Multiple Accountability

Should the duties of public accountability, whether for government as a whole or for particular government agencies, be concentrated in a single person or dispersed among a number of different people? The main argument for concentration lies in the value of having a designated person, usually the leader or agency head, who is obliged to take collective responsibility and to answer to the public, particularly in times of crisis or government failure. Where responsibility and accountability are dispersed between members of a group or between different agencies, officials and politicians can easily shift the blame to others, with the result that no one accepts an obligation to answer to the public. A single point of accountability, by contrast, makes buck passing much more difficult.

On the other hand, multiple avenue of accountability can provide the public with greater opportunities for extracting information from government and for holding government officials up to scrutiny. A leader exercising sole powers of accountability, while less open to buck passing, is better able to resist embarrassing inquiry into the actions of subordinates and to cover up mistakes. In practice, most collective actions involve the responsibility of many individuals ("the problem of many hands"), and the attempt to hold only one person responsible and accountable obscures the reality of how bureaucracies operate.

In general, multiple avenues of accountability appear superior for the initial stages of accountability, for revealing information, and for encouraging the scrutiny of government. Single points of accountability, however, are more effective for taking charge and imposing remedies. The differing U.S. and UK constitutions illustrate the contrast. The United States, with its separation of powers and multiple points of authority, provides a very open and transparent system of government but one where solutions are hard to impose. The United Kingdom, on the other hand, concentrates power in the prime minister and cabinet under tight conventions of ministerial responsibility, which have preserved executive secrecy but allowed effective rectification when problems come to light. A similar contrast can be found in the accountability of federal systems compared with that of unitary systems. Federations provide more avenues for inquiry but lack a single point of responsibility and are prone to blame shifting.

These dilemmas remain intractable. The best solution is to try to combine multiple avenues for scrutiny with a single authority for rectification, thus drawing on the virtues of each approach. Indeed, pluralistic democratic systems of accountability can be viewed in this light. Ideally, they combine many different, complementary mechanisms of scrutiny with a few clear points of unambiguous control and direction.

Accountability in Networks

"Networks" are an increasingly important feature of modern government posing particular problems of accountability. Networks may be understood as structures of collective action and decision making in which formally independent groups or individuals cooperate for shared purposes. They are commonly contrasted both with hierarchies, where the members are linked by formal control structures and in superior–subordinate relationships, and with markets, where self-interested parties are linked through formal agreements and contracts. Network members share the formal independence of market players while cooperating in shared values and objectives over a substantial period of time.

The concept of network has come to prominence in the analysis of modern systems of *governance* (to use a closely associated term) for various reasons. First, it signifies in part the long-standing aspects of all political systems, particularly the more informal cross-institutional relationships, which tended to be overlooked in more traditional institutional analysis and which now appear worthy of much more careful study. For instance, relationships between different levels of government, central and local, federal and state, have always relied heavily on informal partnerships and negotiation. Second, government systems are making more use of arms-length institutions, such as executive agencies and private organizations, for the delivery of public services. Relations between the purchasing government and the provider organizations are neither hierarchical nor market based but instead depend more on network characteristics of partnerships, trust, and agreed values.

Networks cause accountability problems because responsibility for collective action is shared between a number of different parties, giving rise to the classic buck passing associated with multiple accountability. Within the network itself, admittedly, relationships of trust and common interest may lead to mutual accountability and responsiveness. But for outsiders wishing to hold the network to account for its actions, the blurring of responsibility is problematic. Particularly when mistakes are made, interested members of the public are often unable to gain satisfactory answers because the various members of the network can shift the blame to each other.

Outsourcing of public services to stand-alone agencies has been bedeviled by accountability difficulties as politicians blame providers for performance failures and providers respond that the government has given them insufficient resources. The accountability deficit often increases if the providers are in the private sector and not accustomed to the level of public scrutiny applied to the public sector. Commercial contractors plead commercial confidentiality as a reason for concealing their internal operations and are often exempt from investigation by legislative committees or ombudsmen. Nonprofit organizations rely heavily on the conscientiousness of their staff, many of whom are volunteers, and are particularly reluctant to face demands for information and rectification.

Accountability and the New Public Management

The international public sector reform movement of the 1980s and 1990s, known as "the new public management," though primarily aimed at improving public sector efficiency and effectiveness, also included an accountability agenda. The movement's main assumption was that the public sector was less efficient than the commercial private sector and needed to move closer to private sector management methods. With respect to accountability, the most serious deficiency of the public sector compared with the commercial private sector is perceived to be its lack of clear objectives. Private companies have a clear and quantifiable "bottom line," the maximizing of shareholder value, which provides a clear focus for the performance and accountability of managers. Government agencies, by contrast, often have little sense of their objectives and lack clear criteria for judging the performance of officials and their agencies and for holding them to account.

Another deficiency noted in public sector accountability is bureaucrats' comparative lack of concern for serving the members of the public with whom they have direct dealings. In contrast to those offering goods and services for sale in a competitive market—who must focus on their customers' preferences—bureaucrats can exploit their monopoly position and remain largely unaccountable to the members of the public they are supposed to serve.

In other respects, however, the public sector is seen as laboring under excessive accountability burdens. For instance, lack of accountability for results is typically balanced by much more accountability for following set procedures than is found in the private sector. This accountability for process encourages red tape and discourages managerial initiative. At the same time, the managerial efficiency of public sector managers is also stifled by the constant threat of interference in their decisions by their political masters who are themselves responding to the accountability demands of the general public. Because political leaders can be held publicly accountable for any action taken by their departmental officials, the first imperative on all loyal bureaucrats is to save their masters from political embarrassment, even if organizational efficiency is compromised thereby.

The managerial accountability agenda of the NPM reformers therefore includes a number of interlocking strategies. First, objectives are to be clarified, and managers are to be held accountable in terms of achieving measurable outputs. Second, political accountability through elected political leaders is to be confined to the setting of broad objectives and outcomes, leaving responsibility and accountability for outputs with arms-length managers who are quarantined from day-to-day political interference and are accountable to independent regulators rather than to politicians. Third, service providers are to be made more directly accountable to individual citizens, viewed as clients or customers.

The reform agenda has had considerable impact on systems of bureaucratic accountability. Most government agencies now report performance in terms of objectives, including outcomes and outputs, and much effort has been directed toward designing performance measures, particularly in service agencies and in public health and education. The widespread decoupling of service provision from direct political control, through executive agencies and outsourcing, has raised the accountability profile of many agency heads and nongovernment service providers, while removing some internal organizational matters from, public scrutiny. The "audit explosion," signified by the creation of new monitoring and regulatory bodies, reflects the move away from political accountability to independent regulation. Moves toward greater client focus, including the service charter initiatives, pioneered in the UK, have done much to reorient frontline bureaucrats toward a more user-friendly service culture.

At the same time, however, traditional public sector accountability practices have proved much more resistant to change than the reformers had hoped, largely because of entrenched public expectations about accountability. The managerialist injunction to restrict the politicians' accountability for general objectives and to delegate accountability for implementation to agency heads and managers has proved politically unworkable. Members of the public and the media will not readily accept what they see as unjustifiable blame shifting from leaders to subordinates or contractors. Leaders become inevitably drawn into discussion of administrative details: making bureaucrats, correspondingly, remain highly sensitive to political direction. Private contractors become similarly adept at anticipating political pressures if they want their contracts renewed. Hopes that public servants would be less process driven have not been fulfilled. In particular, the continuing role of courts and quasi-judicial tribunals in reviewing administrative decisions has maintained a strong demand for due process. The public sector continues to be held to higher procedural standards through the principles of natural justice, which are jealously protected by the courts.

Richard Mulgan
Australian National University
Canberra, Australia

See also Audit Society; Auditing; Bureaucracy; Contract Theory; Networks; New Public Management; Performance Management; Principal–Agent Theory; Regulation; Responsibility; Responsiveness; Rule of Law

Further Readings

Behn, R. (2001). *Rethinking democratic accountability*. Washington, DC: Brookings Institution.

Bovens, M. (1998). *The quest for responsibility: Accountability and citizenship in complex organisations*. Cambridge, UK: Cambridge University Press.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal, 13,* 447–468.

Day, P., & Klein, R. (1987). *Accountabilities: Five public services*. London: Tavistock.

Finer, H. (1941). Administrative responsibility and democratic government. *Public Administration Review, 1,* 335–350.

Flinders, M. (2002). *The politics of accountability in the modern state*. Aldershot, UK: Ashgate.

Friedrich, C. J. (1940). Public policy and the nature of administrative responsibility. In C. J. Friedrich & E. S. Mason (Eds.), *Public policy* (pp. 3–24). Cambridge, MA: Harvard University Press.

Grant, R. W., & Keohane, R. O. (2005). Accountability and abuses of world power. *American Political Science Review, 99,* 29–43.

Harlow, C. (2002). *Accountability in the European Union*. Oxford, UK: Oxford University Press.

Keohane, R. O. (2003). Global governance and democratic accountability. In D. Held & M. Koewnig-Archibugi (Eds.), *Taming globalization: Frontiers of governance* (pp. 130–157). Cambridge, UK: Polity.

Mulgan, R. (2000). "Accountability": An ever-expanding concept? *Public Administration, 78,* 555–574.

Mulgan, R. (2003). *Holding power to account: Accountability in modern democracies*. Basingstoke, UK: Palgrave Macmillan.

O'Donnell, G. (1998). Horizontal accountability in new democracies. *Journal of Democracy, 9,* 112–126.

Philp, M. (2009). Delimiting democratic accountability. *Political Studies, 57,* 28–53.

Przeworski, A., Stokes, S. C., & Manin, B. (Eds.). (1999). *Democracy, accountability and representation*. Cambridge, UK: Cambridge University Press.

Radin, B. (2002). *The accountable juggler: The art of leadership in a federal agency*. Washington, DC: CQ Press.

Rhodes, R. A. W. (1997). *Understanding governance: Public networks, governance, reflexivity and accountability*. Buckingham, UK: Open University Press.

Romzek, B. S., & Dubnick, M. J. (1987). Accountability in the public sector: Lessons from the Challenger tragedy. *Public Administration Review, 47,* 227–238.

Schedler, A. (1999). Conceptualizing accountability. In A. Schedler, L. Diamond, & M. F. Plattner (Eds.), *The self-restraining state* (pp. 13–28). Boulder, CO: Lynne Rienner.

Accountability, Electoral

Electoral accountability is best interpreted as a three-phase process and a feedback. Political actors looking for election will first try to take into account the preferences of the voters. If elected, they will then keep into account what they have heard and learned during the electoral campaign. Returning to the voters, they will give an account of their behavior and their performance. Even those representatives who do not run again will retain some electoral accountability, because none of them would want to ruin the electoral chances of his or her party and successor.

Electoral accountability lies at the heart of all processes of democratic representation. It is a complex, multilayered phenomenon entailing three quite distinct, though interrelated, phases and several individual and collective actors. It is significantly affected and, in some ways, shaped by the electoral system and, more generally, by the institutional arrangements.

Accountability During the Electoral Campaign

Electoral accountability begins when a relationship is established between the voters and their potential representatives, candidates, and parties. During the electoral campaign, the candidates and the parties have an interest in getting to know what the voters need and want, which preferences and values they have, and which ideals they would promote. In this phase, accountability manifests itself as the most conscious effort by candidates and parties to learn, that is, to *take into account* what the voters communicate to them. Obviously, there are two limits to this process. First, the voters may not have a very precise idea about their preferences and interests and may not know exactly how to communicate them to the candidates and the parties. Second, the candidates and the parties may discount the preferences and the values of some or several groups of voters or may be unable to introduce them into their predisposed programmatic package and, to a much lesser extent, to their ideological views. On the whole, however, an effort will certainly be made by them to take into account what they have heard and have come to know.

Usually, a difference is drawn between a proportional electoral system used in relatively large constituencies with no possibility of casting a preference vote and plurality-majority electoral systems applied in single-member constituencies. Understandably, in the first case, much or most of the electoral accountability depends on the ability and the willingness of the political parties to take into account what they have heard during the electoral campaign, and much of the electoral accountability will then be projected to the national level. When plurality-majority systems operate in single-member constituencies, individual candidates are the protagonists of electoral accountability (Bruce Cain, John Ferejohn, & Morris Fiorina, 1987). It is up to them to interact with the voters, to learn as well as to explain, to take into account what they have been communicated by the voters, and also to carry that knowledge into the representative assembly. The candidates' ability to learn about the preferences of the voters and to shape them may make a difference in the outcome of the electoral processes and, possibly, in the type of politics and policies proposed and later implemented by their party or governmental coalition.

Accountability in the Implementation of Policies

Electoral accountability does not end with the holding of the elections and their outcomes. Once installed inside the representative assembly at any level of the political system, elected representatives and their parties must try to translate their proposals into policies. Again, there is a difference in the approach taken by representatives elected through a proportional system in large constituencies, even more so when there is no preference vote, and representatives elected in single-member constituencies through a plurality-majority electoral system. In the first case, party leaders will decide which policies to implement, usually claiming to have received a mandate and remaining somewhat indifferent to positions articulated by their individual representatives. Party leaders will *keep into account* what they believe has been the response of the voters to their overall programmatic offer. As to representatives elected in single-member constituencies, no doubt quite a number of them will certainly support their party's policies. However, not a small number of those elected representatives will indeed attempt to introduce into the political discourse and the policy debate what they have learned from the/their voters. In addition to some general preferences and interests expressed by their voters and largely compatible with their party's program/manifesto, elected representatives will seriously attempt to keep into account specific preferences and interests concerning their single-member constituencies and their voters, or all of them. Indeed, the working of a specific representative assembly will be significantly affected by the degree of political discretionality and electoral power of its individual members, especially those elected in single-member constituencies (Heinz Eulau, 1986).

Accountability and Reelection

Neither political debates nor electoral accountability comes to an end in a representative assembly. All party leaders and elected representatives, governmental office holders, and oppositionists are fully aware that public opinion has its (more or less open and informed) eyes on them, that it scrutinizes their behavior, its conformity to the promises and the programs, and its consequences. One way or another, all the protagonists, voters included, know perfectly well that in democratic regimes the voters will periodically pass judgment on the policies that they have approved or opposed and on their behavior while in office. Hence, when the time comes for new elections, elected representatives and party leaders cannot escape from performing another task fully belonging to accountability. They will have to explain to the voters what they have done, not done, or poorly done. They will be obliged to *give an account* of their performance in their roles. It is of decisive importance that elected representatives and party leaders accept full responsibility for their performance or lack of it. Indeed, often, but inadequately, electoral accountability is considered mechanically limited only to this specific phase: reward (reelection) and punishment (defeat). No doubt, depending on the strength and the independence of the mass media and on the degree of vibrancy and robustness of civil society, the politicians in office aiming at reacquiring their positions will be held accountable. Technically, the process through which representatives and party leaders launch their reelection bid is called feedback. It introduces old and new issues, renewing the circuit of accountability that depends on what was done and has to be reformed, and what was set aside but must be taken again into account.

Seen from this perspective—that is, focused on the activities, the perceptions, the performance of parties as well as individual representatives—electoral accountability is a never-ending game. However, it is true that from time to time individual representatives do voluntarily retire. It is also true that in some political systems for some political offices, there are term limits that oblige those office holders to exit from politics. One may conjecture that "lame duck" politicians will feel less inclined to take electoral accountability into account in their behavior. They will not be asked by the voters to explain their behavior. Hence, at least in theory, they will run few risks if and when they behave in an irresponsible way. What is missing in all the statements concerning the potential irresponsibility of office holders who are not constrained by the imperatives of reelection is the relationship between those representatives and their corresponding parties. Though there is no specific research on the motivations of the outgoing representatives or on the ability of party leaders to enforce on them the "ethics" of accountability, it does not seem farfetched to suggest that something of the kind is and always has been at work in most cases. That is, practically no outgoing representative has deliberately chosen to behave in an indifferent manner—more precisely, not caring at all about the relationship between his or her promises and his or her behavior, thus negatively affecting the chances of election of his or her party's candidates.

Accountability in Proportional Versus Majoritarian Electoral Systems

A larger issue looms with regard to electoral accountability, though to be more precise, perhaps one ought to speak of political representation. Though the comparative study by G. Bingham Powell (2000) is excellent, unfortunately it does not tackle the various processes, phases, and interpretations of electoral accountability. Rather, it refers to what kind of representation is overall provided by proportional systems (in which a party's share of seats is determined by its share of votes) versus majoritarian electoral systems (in which a candidate must win a plurality or an absolute majority of votes to be elected). Assemblies elected through proportional electoral systems and coalition governments may provide for the representation of a wider spectrum of opinions, interests, and preferences. Using the terminology above, they have the possibility of taking into account a greater number of opinions, interests, and preferences in the first phase of the electoral process. Statically, they may almost come to "mirror" them. Not much is learned, however, when it comes to the phase in which those opinions, interests, and preferences will have to be kept into account. The representatives of the various parties and their leaders will attempt to have their way, and the ensuing bargaining process does not at all guarantee better accountability, in terms of transparent and responsible processes of decision making. Also, when returning to the voters to give them an account of what has been done or not done, it is likely that the politics of *buck passing* will defeat the politics of personal and party accountability. It is also likely that some representatives and several party leaders might enact the politics of *outbidding* if they believe that they will not be asked to take on governing duties.

On the contrary, in single-member constituencies and in one-party governments, one can legitimately hypothesize that two likely processes will be at work. All representatives elected in single-member constituencies will feel the need to avoid relying exclusively on their "initial" voters. Some of those voters will inevitably be dissatisfied with the performance of their specific representative. Thus, the representative will attempt to increase the number of his or her voters by taking into account a larger set of preferences and by keeping them into account in his or her representational and governing activities in order to be reelected. A similar process will take place where there is single-party government. In their governing activities, all one-party governments are bound to lose some of their initial supporters. Hence, party leaders will launch their representational net beyond the initial borders of their electoral consensus. By so doing, they will attempt to take into account more interests, more preferences, and more opinions than simply those of the people who had voted their party into office. The process of widening the perimeter of their support seems much more difficult for representatives elected on party lines and for multiparty governments. Each party representative will have to toe the party line lest he or she jeopardizes his or her reselection. Hence, they will all stick to a static interpretation of their accountability: taking and keeping into account exclusively what was learned and promised through party channels. No party belonging to the governmental coalition will be likely to afford looking for additional outside support lest the political jealousy and envy of the other parties making up the governmental coalition lead to its breakup.

All these, admittedly, though indispensably, hypothetical, considerations are not necessarily meant to prove that majoritarian electoral systems applied in single-member constituencies are absolutely preferable to proportional electoral systems from the point of view of accountability, not the reverse. What they suggest is that no final conclusion concerning the superiority of proportional systems in terms of accountability seems justified. Only empirical research done countrywide and based on different electoral systems will offer sufficient and satisfactory material to understand and assess the quantity and quality of electoral accountability.

Gianfranco Pasquino
University of Bologna
Bologna, Italy

See also Accountability; Accountability, Interinstitutional; Representation; Responsibility; Responsiveness

Further Readings

Cain, B., Ferejohn, J., & Fiorina, M. (1987). *The personal vote: Constituency service and electoral independence*. Cambridge, MA: Harvard University Press.

Eulau, H. (1986). *Politics, self, and society*. Cambridge, MA: Harvard University Press.

Powell, G. B., Jr. (2000). *Elections as instruments of democracy: Majoritarian and proportional visions*. New Haven, CT: Yale University Press.

Accountability, Interinstitutional

Interinstitutional accountability is a form of horizontal accountability in which the different institutions provide a check on one another's behavior. The legislature is supposed to check the actions of the executive branch of government. The judiciary, especially constitutional courts, oversees the behavior of governing and representative institutions. Independent authorities—that is, all those agencies with power of oversight on specifically relevant aspects, such as central banks, regulatory agencies, and ombudspersons, or those authorities who oversee the regulation of media and communication or of competition—indirectly check the government and can be controlled by it in their actual working. Party government remains the best "mechanism" for interinstitutional accountability when party appointees and party representatives are accountable horizontally to their leaders and vertically to the voters.

This entry first discusses the definition of interinstitutional accountability and its evolution in Europe and the United States. It then considers ways in which such accountability has been expanded in recent decades to include the roles of institutions such as central banks, regulatory agencies, and ombudspersons.

The Nature of Interinstitutional Accountability

Institutions may be accountable vertically and horizontally. *Vertical* accountability, that is, electoral accountability, exists when institutions are exposed to the control of the voters and the citizens. Technically, the voters and the citizens are the principals and the institutions are the agents.

The citizen/voters delegate some representatives in institutions such as government and parliament to perform some tasks. They retain the capacity and the power to reward or punish a government and/or a parliament/congress. In this situation, vertical accountability is said to be at work. This type of accountability is the hallmark of democratic regimes. Where vertical accountability does not exist, there is no democracy, though, of course, a democratic regime requires and provides much more than just "vertical accountability." On the other hand, *horizontal* accountability is a relationship established among institutions: the executive, the legislative (parliament or congress), the bureaucracy, the military organization, the judiciary, and independent authorities. It occurs when the various institutions in different ways are in a position to control the activities of another specific institution. This type of control can be reciprocal and balanced, one-sided and skewed, formal and informal, accompanied by predefined sanctions, or largely symbolic.

Notwithstanding their often very significant differences, all these relationships can be legitimately assigned to the area of what is called "interinstitutional accountability." In a way, horizontal-interinstitutional accountability is a much more complex relationship to be studied, first of all, because it is not dyadic (voters–leaders) and may entail the activities of more than two institutions. Second, not only is it less transparent than any electoral process, but it is also less streamlined. Interinstitutional accountability is also made more complicated by the sheer fact that it is the product of a fair amount of reciprocal interaction and mutual influence. For this reason, its full understanding requires a profound knowledge of the dynamics of the entire political-institutional system. Finally, the study of horizontal-interinstitutional accountability is also more likely to be influenced by the values of the scholars, who may prefer the stability to the instability of the policy process and the concentration of political decision-making power to its diffusion, or, more rarely, vice versa. In different periods of time, some of these preferences, and values, advocated and shared by scholars and power holders, have been translated into concrete institutional arrangements. Indeed, a short exploration of selected constitutional outcomes is a good introduction to the analysis of interinstitutional accountability.

According to Aristotle, mixed regimes are the preferred form of government. If one political actor prevails over all the others, it results in tyranny. If few actors prevail over all the others, the regime is an oligarchy. When many actors have unbridled power, the regime runs the risk of degenerating into what Aristotle called democracy, the positive definition of a well-functioning and balanced regime being *politeia*. When power is highly concentrated in the hands of just one, a few, or many actors, it is impossible to control the exercise of that power and to impose accountability on the power holders. On the contrary, when it became possible to allocate power to different institutions, through the appearance, for instance, of constitutional monarchies, the seeds of interinstitutional accountability were sown.

Evolution of Interinstitutional Accountability

Most scholars would concur that the famous book by Charles-Louis de Secondat, Baron de La Brède et de Montesquieu, *De l'esprit des lois* (1748), represents the first attempt to suggest the division of institutional power, though not yet the creation of mechanisms capable of producing interinstitutional accountability. At the time, the problem to be solved was how to take away from the king the power to legislate and the power to adjudicate and give them to two specialized institutions: parliament and the judiciary. In practice, the king retained some legislative and judicial powers, but the very existence of two institutions enjoying most of those powers meant the first appearance of some relatively simple tripartite interinstitutional accountability. The separation of institutional powers created a situation in which different institutions were in a position to control each other, giving birth to the circuit of accountability among the three most important institutions.

The next, highly significant and quite deliberate, attempt to construct a web of interinstitutional accountability—the drafting of the Constitution of the Republic of the United States of America—constituted a response to practical problems more than a full-blown theoretical formulation. In all likelihood, James Madison and Alexander Hamilton thought that they were going to substantially imitate the institutional edifice existing in Great Britain. But they were also aware of the need to decisively innovate and improve on it. Montesquieu had fundamentally aimed at subtracting the legislative and the judicial power from the executive power (the king). The three powers could then act autonomously and independently, giving birth to a situation of positive institutional pluralism. However, neither autonomy nor independence was in itself leading to interinstitutional accountability. On the contrary, in Montesquieu's view, the institutional *separation* of the three "powers" was positive in itself. It was on the other side of the English Channel that a somewhat different interinstitutional formula made its appearance: "the king in parliament," meaning a close relationship and some reciprocal accountability between the chief executive and the representative assembly. The king could prevent parliament from passing undesired laws, and in its turn, parliament could oppose bills and, above all, taxes desired by the king. The product of a prolonged and even bloody confrontation, this apparently healthy situation represented a most prominent instance of "checks and balances."

At the time the U.S. Founding Fathers were drafting, defending, and promoting their constitutional package, Madison and Hamilton shared the view that in the English case one could find some separation of the institutions combined with some interaction among them. Both of them, but especially Madison, argued that the separation of the institutions was a good formula to prevent the accumulation and concentration of power in one specific institution. Though Hamilton wanted more power to be given to the presidency, in the end Congress, being more representative of the American people, received more and more important powers. And the U.S. political system parted forever from the institutional architecture of the British constitutional monarchy.

Great Britain

The development of the political and institutional dynamics of Britain led to the *fusion of powers*, the secret of the English Constitution's efficiency, according to Walter Bagehot (1867). That is, the relationship between the English cabinet and parliament is clearly skewed in favor of the prime minister, who obtains and retains his or her office because he or she is the leader of the parliamentary party that owns the absolute majority of seats. As Max Weber correctly wrote, the English prime minister is "the dictator of the parliamentary battle field." Hence, the separation does not run neatly between the prime minister with the cabinet and the parliament, between 10 Downing Street and Westminster, but between the prime minister together with his or her parliamentary majority and the opposition led by its shadow government. In a way, Winston Churchill was wrong: It is not parliament that is fully sovereign and omnipotent, but the parliamentary majority led by its prime minister. As a consequence, the English cabinet is accountable not to parliament as such but above all to its own parliamentary majority. The relationship of accountability does not occur between, on one side, government and, on the other, parliament, but between the governmental majority and the opposition. Still, under some exceptional circumstances, when party discipline breaks down, parliament may and will revive the circuit of interinstitutional accountability with Her Majesty's government. A similar situation holds in those parliamentary democracies that have imported the British model, for example, Australia, Canada, and, until 1993, New Zealand, plus some former British colonies in the Caribbean and Africa. In all these cases, an effective two-party system seems to be the necessary condition for interinstitutional accountability following the English style.

United States

Notwithstanding the oscillations between congressional government and the Imperial presidency, the definition that best characterizes the functioning of the U.S. political system was formulated by Richard Neustadt (1960): "separate institutions sharing powers." This widely accepted interpretation depicts a reality that is quite different from, one might even say almost the opposite of, the English institutional arrangement. In the United States, the executive and the legislative powers, the president and Congress, are the product of distinct electoral processes. The president does not have the power to dissolve Congress, while Congress cannot dismiss the president except through the very extraordinary procedure of impeachment, which, incidentally, does represent an extreme instance of interinstitutional accountability. Congress and the president are obliged to live and work together lest a paralysis of the political system follows. They are both endowed with quite powerful instruments to check their respective activities and behavior that, in some exceptional instances, for example, when they disagree on the federal budget, may lead to decision-making paralysis or a stalemate.

Including the judiciary, especially, the Supreme Court, in the overall framework, the situation of "separate institutions sharing powers" is complete. The president holds the power to appoint the justices (as well as several other federal judges), but the Senate retains the power to reject those appointments ("advice and consent"). The Senate has even the power to block the appointment of all the secretaries nominated by the president. In both instances, interinstitutional accountability may also be at work in a rather subtle way: deterrence and anticipated reactions. Fearing the opposition of the Senate, the president will avoid nominating someone who is likely to be rejected. On its part, the Senate will refrain from challenging the president on all nominations. Once appointed, the Supreme Court justices will hold office for life, which allows them to behave with absolute freedom, not being subject to any kind of retaliation. In a way, their accountability stops, but they are in a position to impose a lot of accountability both on Congress and on the president. The members of the European constitutional courts are not appointed for life but for fixed terms. Many of them have a career of their own after their experience in the constitutional court. It is not rare for them, as in the case of Italy, to position themselves, in some cases at the expense of interinstitutional accountability, to obtain other offices.

As to the interinstitutional accountability occurring between the U.S. president and Congress, though Congress enjoys the power to initiate legislation, it is exposed to presidential vetoes, which are often threatened and applied and which, for want of a three-fifth majority, are rarely overridden. A president's agenda can generally be achieved only in those infrequent cases in which the president's party has majorities in both branches of Congress that are solid and disciplined. Even the formation of the presidential team of secretaries, the so-called administration, needs the approval of—that is, confirmation by—the Senate. All these complex processes justify Neustadt's definition and evaluation. Indeed, U.S. institutions, though separate, share powers and compete for the exercise of those powers, making space for many opportunities of interinstitutional accountability. This is even truer when the president's party does not have a majority in one or both branches of Congress. A *divided government* may entail not just more interinstitutional accountability but also political and institutional confusion, stalemate in decision making, and pork barrel policies. It is not at all farfetched to hypothesize that confusion in the interinstitutional accountability process will negatively affect the possibility of vertical accountability as well. A number of principals (the citizen-voters) will find it difficult to identify which agents (the president or the members of Congress) to punish or reward. At this point, a classic institutional problem emerges: How much are we willing to sacrifice in terms of decision-making efficacy in order to obtain and maintain democratic control?

France

The quite different institutional arrangement of the Fifth French Republic, and more generally of semipresidential systems, entails different consequences in terms of interinstitutional accountability. In the French case, interinstitutional accountability is a more complicated phenomenon because it depends also on the changing allocation of political power between the president and the Assemblée Nationale (and, eventually, the leader of its parliamentary majority). Reacting against the institutional weakness of the Fourth Republic, Charles De Gaulle favored the concentration of much power in the hands of the president and the president's government. Semipresidentialism was meant to give more, and, in a way, "shielded" power to the chief executive, in order to endow the chief executive with a lot of (popular) legitimacy, certainly not to establish a circuit of interinstitutional accountability. De Gaulle never envisaged the likelihood of *cohabitation*, that is, of a president having to live with a party of the opposite political alignment that has the parliamentary majority and a prime minister who belongs to that party. Cohabitation introduces a significant amount of interinstitutional accountability, though at a price. Unlike the U.S. president (and all chief executives in presidential systems), the French president has the option of dissolving parliament in the hope that the voters will return the majority he or she likes. The president can dissolve parliament only after it has been in office for at least a year. However, no president can risk tarnishing his or her image by having the party with the opposite political alignment repeatedly returned by the voters. Hence, the network of interinstitutional accountability is enriched by the "interference" or the calling of the voters into the picture rather than by staying within the limits of interinstitutional relationships. A prime minister who belongs to a political alignment opposed to the president will avoid deliberately challenging the president, because such a move would risk the early dissolution of a parliament in which the prime minister's party is in the majority. This potential for the dissolution of parliament thus obliges both the president and the prime minister to behave accountably and, on the whole, predictably. Though both may engage in *buck passing*, the citizen-voters are in a position to evaluate specific responsibilities and to punish either the president for exaggeratedly interfering or the prime minister and his or her majority for their (lack of) performance. Finally, there is another important structure enforcing some interinstitutional accountability: the Constitutional Council, which can be activated by just 10% of the members of the Assemblée Nationale to challenge the bills approved by the majority.

Germany

All parliamentary governments are built on more or less complex and explicit webs of interinstitutional accountability. All governments must have and maintain a relationship of confidence with parliament, otherwise they will be replaced more or less smoothly. The German constructive vote of no confidence represents a specific mechanism of this type of accountability. A parliament that is unable to support its government or unwilling to enact its policies can be dissolved either by the head of the government or by the head of the state following a request by the head of the government.

German constitution makers also wanted to prevent the concentration of power. The semifederal system, based on the autonomy of the *Länder*, was the answer, especially strengthened by the many and important legislative and representative powers given to the second chamber: the *Bundesrat*. Though not involved in the procedure of voting confidence in the chancellor and withdrawing it, the Bundesrat is an important actor in the circuit of interinstitutional accountability. It has significant powers in all legislation concerning even marginally the policies considered in the domain of the Länder. Its power to prevent the passing of legislation approved by the *Bundestag* compels this assembly to be accountable. More generally, all types of federalism and all forms of political and administrative decentralization translate themselves into some interinstitutional accountability between the federal government and local governments.

Emerging Roles of Other Agents of Accountability

When speaking of interinstitutional accountability one must consider another, relatively recent phenomenon that identifies other authorities and agencies that can be instruments of accountability, such as central banks, regulatory agencies, and ombudspersons.

First, many countries have granted full independence to their central banks. The U.S. Federal Reserve has enjoyed such independence for some time, while the British Central Bank acquired it in 1997. With several other countries following suit, this trend means that one can find many instances in which interinstitutional accountability between the treasury and the central banks is at work.

Second, several countries have created independent authorities and agencies charged with tasks once performed by the executive or, at least, supposed to be in its domain or that of the legislature. In the second case, these agencies may complement legislative oversight. Examples are agencies that address areas such as consumer protection, food and drug safety, communications, and privacy issues raised by the news media; enforce regulations intended to ensure a competitive market; and oversee evolving energy technologies. In many cases, those agencies and authorities, whose chairs and boards are appointed either by government or by parliament, report directly to parliament.

Finally, in many countries and even in the European Union (EU) there is an ombudsperson who deals with the complaints of the citizens against their state and its bureaucracy as well as the very decisions of the institutions of the EU itself. Also, within the EU, interinstitutional accountability is present in its complex circuit going from the European Council to the European Commission to the European Parliament, but it also plays a role in the relationships between the institutions of the EU and the member states.

It is unclear whether the proliferation of authorities and agencies has actually improved the working and quality of interinstitutional accountability. When dealing with complex processes whose components, antecedents, and consequences are not and cannot be crystal clear, no attempt at evaluation is bound to be satisfactory. One may want to point to the perils of too dense and wide a network of interinstitutional accountability. The first peril has certainly to do with the (im)precise definition of the spheres of action of several institutions. In too many instances, tensions and conflicts may appear between the executive and the legislative, between the federal government and local governments, or, quite often, between the executive and the judiciary. It will then be the constitutional court that decides where the respective spheres of action begin and, above all, end. But there will always remain a gray area. Some institutional actors will try to take advantage of this gray area and engage in the politics of buck passing, that is, attempting to put the blame for what is done, not done, or poorly done on other institutions. In addition to tensions and conflicts, confusion will follow. The perils of interinstitutional accountability will affect the democratic framework itself. Though by definition the citizens are not directly involved in this kind of accountability, still they are more or less informed spectators who can look at what goes on and who can form their own opinion. Not being directly in the position to provide rewards or punishments, they are likely to reach a negative judgment on all the institutions involved in this exercise of political and bureaucratic buck passing. If and when all the institutions are considered equally responsible for the conflict and for the confusion, some political apathy or even alienation cannot be written off.

Technically, the mass media are not considered to be fully a component of the network of interinstitutional accountability. Nevertheless, they are often called the fourth power. But are they accountable to other institutions? Leaving aside their somewhat peculiar accountability to their readers and public opinion at large, at the most the mass media are accountable to the law, hence indirectly to the judiciary. They do play a political role that most certainly affects the way interinstitutional accountability is performed and evaluated, but the mass media themselves are external to that circuit.

Conclusion

In sum, there is a vibrant paradox in all discussions of interinstitutional accountability. Most scholars have maintained that this kind of accountability is quite positive for the working of democratic regimes. It seems to be less positive for the power of the people. Political power should not be concentrated but distributed among several institutions and shared by them. Checks and balances must be devised in such a way as to put a limit on the exercise of power by all institutions. Some activities are so important that they have to be controlled by special bodies: independent authorities and agencies. Recently, doubts have been raised concerning not only the effectiveness of interinstitutional accountability but also, inevitably, whether a dense network of interinstitutional accountability may in practice conspicuously trim all the occurrences of political accountability in which the citizens are the principals. When accountability escapes evaluation by the citizens because some policy areas are insulated and not exposed to their control, the quality of democracy may significantly suffer. Some nostalgia has made its appearance for the period and the situations in which the *party government* and a partisan opposition faced each other in a visible and transparent competition decided by the voters. There is more than a kernel of truth in the statement that such competition in contexts in which governing parties do not delegate power to outside agencies and authorities does provide for better accountability. When the parties in government are many, the voters will still encounter problems in their attempt to punish and reward their "agents." But most of the time it will be clear that parties and their leaders are accountable for the positive and negative functioning of the governing, representative, and bureaucratic institutions.

The extension of instances of interinstitutional accountability may have been fueled and nourished by distrust toward parties and party politicians. It may have had as a goal the containment and the reduction of party power. But in many cases, politicians seem to have reacquired their power working within and through the institutions, thus avoiding electoral accountability. When necessary, party leaders have learned to practice and implement a shrewd politics of appointments of loyal, even subservient, collaborators, in all nongovernmental bodies and agencies, guiding them to enact unpopular policies. The proliferation of institutions endowed with some decision-making power in technical areas or with the power to check the activities of other institutions has seemed to some a democratic conquest, to others an expropriation of the voters. Most of the time, it is still party politicians who control the exercise of political power by and within the institutions, while it is appointed technocrats and career bureaucrats who come to bear the brunt of the blame in the interinstitutional web of accountability. Not all that glitters is gold (democratic) in the circuit of interinstitutional accountability.

Gianfranco Pasquino
University of Bologna
Bologna, Italy

See also Accountability; Accountability, Electoral; Representation; Responsibility; Responsiveness

Further Readings

Bagehot, W. (1867). *The English constitution*. Oxford, UK: Oxford University Press.

Lewin, L. (2007). *Democratic accountability: Why choice in politics is both possible and necessary*. Cambridge, MA: Harvard University Press.

Neustadt, R. E. (1960). *Presidential power and the modern presidency*. New York: Free Press.

Przeworski, A., Stokes, S. C., & Manin, B. (Eds.). (1999). *Democracy, accountability, and representation*. Cambridge, UK: Cambridge University Press.

Strøm, K., Müller, W. C., & Bergman, T. (Eds.). (2003). *Delegation and accountability in parliamentary democracies*. Oxford, UK: Oxford University Press.

ADMINISTRATION

Governments in many Western countries decided in the 1980s that traditional administration was no longer up to the task of modern government. They looked to the private sector to define a new management approach. Management introduces a new vocabulary, mindset, and culture to government bureaucracies. The purpose is to force the hand of bureaucrats to become more dynamic and better managers. A debate has raged in the political science and public administration literature on the merits of the new public management (NPM) in relation to traditional public administration for the past 25 years. A number of politicians in the 1980s decided that the machinery of government was in an urgent need of repair. The old ways were no longer up to the task, and they packaged a series of reform measures that in time would become known as NPM. Proponents of NPM deliberately set high standards with goals to "reinvent government," "get government right," and designate those that would "steer" government from those that would be doing the "rowing."

The word *management* in NPM implies a decisiveness, a bias for action, and a dynamic mindset. Traditional public administration, meanwhile, conjures up images of rules, regulations, and lethargic decision-making processes. Presidents and prime ministers who came to power in the 1980s concluded that the problem was with bureaucracy, not political institutions. They accused bureaucracy of being bloated, expensive, unresponsive, a creation of routine deliberately resistant to change, and essentially incapable of dealing with new challenges.

It is one thing to diagnose the patient, but it is quite another to come up with the remedy. Initially, at least, political leaders were left to try this or that to see what would work with varying degrees of success. In time, a new approach, anchored in private sector management practices, began to take shape and a label was attached to it—NPM.

The goal was nothing short of introducing a new culture in government departments and agencies. The old culture was found wanting on many fronts. It attached too much importance to due process, prudence, probity, and centrally prescribed administrative rules and regulations. It also encouraged senior civil servants to focus on policy issues rather than on management. The old culture, associated with traditional public administration, was considered not only outdated but also counterproductive. Public administration became synonymous with the old culture.

NPM would introduce a new vocabulary, a new way of thinking, and a new culture. It would also give rise to a proliferation of management techniques to force government operations to become more efficient. The purpose was to force the hand of senior government officials to become better managers and to learn to make tough management decisions. Taken at face value, the political rhetoric that accompanied the arrival of NPM to government meant setting the civil service at its own throat.

The Canadian government published a report (see Table 1) designed to contrast the old culture (public administration) with the new culture (NPM).

NPM holds important advantages for politicians. In forcing the hand of senior civil servants to become better managers, politicians would gain the upper hand in shaping policy initiatives. The thinking was that civil servants had too much influence on policy at the expense of politicians. The message from politicians to senior civil servants could not be clearer under NPM—you worry about managing government operations better, and we will worry about setting policy priorities.

Table 1 Old Culture (Traditional Public Administration) Versus New Culture (New Public Management)

Old Culture	*New Culture*
Controlling	Empowering
Rigid	Flexible
Suspicious	Trusting
Administrative	Managerial
Secret	Open
Power based	Task based
Input/process oriented	Results oriented
Preprogrammed and repetitive	Capable of purposeful action
Risk averse	Willing to take intelligent risks
Mandatory	Optional
Communicating poorly	Communicating well
Centralized	Decentralized
Uniform	Diverse
Stifling creativity	Encouraging innovation
Reactive	Proactive

Source: Government of Canada, Public Service 2000 Secretariat, Ottawa, Canada.

How then would civil servants become better managers? NPM encompasses a number of broad strategies to promote cultural change in government: decentralization of decision making, empowerment, a reduction on controls on managers, more flexible organizational structures, upgrading the skills of government managers, and a stronger sense of service to the public. Government departments and agencies were also encouraged to launch review exercises to identify "useless" red tape and "delayer" management levels.

In time, NPM became the fashion in much of the Western world. Margaret Thatcher in the United Kingdom (UK) showed the way with numerous measures designed to overhaul government operations. She cut the size of the civil service (from 733,000 strong to 569,000), restructured government operations by creating executive agencies and gave them a narrow mandate to deliver public services, privatized state corporations, delegated more authority to frontline managers, and overhauled the government's financial management system. This and other private sector management–inspired measures gave life to NPM in the UK. Before long, Australia, New Zealand, the United States, and Canada, among others, also introduced numerous NPM-type measures with varying degrees of success. Countries that did not pursue NPM with any enthusiasm, such as France, were regarded as being out of step with modern management strategies.

The contrast between the old (public administration) and the new (NPM) is striking and instructive. However, it probably came as a surprise for retired civil servants to discover that their culture was "rigid" and gave rise to "suspicious" and "secretive" behavior, which in turn served to "stifle creativity," which led them to "communicate poorly."

Old cultures die hard, if they die at all. NPM has met some successes but also failures. Graham Allison (1987) went to the heart of the matter when he wrote,

> The perception that government performance lags behind private business performance is

> correct. But the notion that there is any significant body of private management practices and skills that can be transferred directly to public management tasks in a way that produces significant improvements is wrong. (p. 525)

The sharp differences between the public and private sectors make it difficult for a number of NPM measures. The public sector does not have a clear bottom line and has legislative requirements of reviewing the work of government managers. Indeed, public sector managers depend on political bodies both for their authority and for their budgets, and they are subject to far greater public scrutiny of their actions. They are directly accountable to their political masters, and no matter what form of government organization is in place it can only enjoy legitimacy through the political process. Although it can take various forms, governments must have an accountability process to make the exercise of power responsible. This speaks of the merits of traditional public administration.

Traditional public administration has shortcomings, but it is directly tied to national political institutions and their accountability requirements. Hierarchy, command and control, and centrally prescribed rules and regulations enable political leaders to reach down to all departments and agencies and determine who did what and why things went wrong.

NPM had to compensate somehow for its call to empower managers, to flatten organizations, and to import numerous private sector management practices to government. An emphasis on measuring performance became another defining characteristic of NPM. Nothing is left in government that is not up for measurement from the performance of senior civil servants to program activities in all sectors whether economic or social. The thinking is that establishing performance standards will provide for more effective accountability and also enable both politicians and citizens to see how well civil servants and programs perform.

The demand for information to fuel performance measurement and evaluation initiatives by central agency and senior departmental officials has increased substantially in recent years throughout the Western world. It seems that everything in government now needs to be measured to have any standing in the expenditure budget process. This squares with NPM's emphasis on outputs, on good management, and on evaluating how well individual civil servants and programs perform. This too is in contrast to traditional public administration with its emphasis on controlling input costs and holding civil servants accountable for their administrative decisions.

Proponents of traditional public administration insist that politics and the public sector do not lend themselves to big answers. Public administration operates in a political environment that is always on the lookout for "errors" and that exhibits an extremely low tolerance for mistakes. The attention of the national media and political opponents are sufficient to explain why civil servants are cautious and why they strive to operate in an error-free environment. One would have to let the imagination run wild to visualize a headline in the media applauding the fine work of "empowered" civil servants. Supporters of traditional public administration insist that it is unwise to think that one can import private sector management strategies to government. They point out that in business it does not much matter if you get it wrong 10% of the time as long as you turn a profit at the end of the year. In government, it does not much matter if you get it right 90% of the time because the focus will be on the 10% of the time you get it wrong.

In the end, the debate that truly matters between the two camps centers on accountability. Both sides insist that their approach holds greater merit in ensuring that politicians and civil servants are accountable for policies and the delivery of government programs. For politicians and civil servants, both approaches have strengths and drawbacks—traditional public administration provides more stability and predictability while NPM provides greater flexibility and a capacity to measure the performance of both senior civil servants and programs.

Donald J. Savoie
Université de Moncton
Moncton, New Brunswick, Canada

See also Administration Theory; Autonomy, Administrative; Civil Service; Development Administration; Performance Management

Further Readings

Allison, G. T. (1987). Public and private management: Are they fundamentally alike in all unimportant respects? In J. M. Shafritz & A. C. Hyde (Eds.), *Classics of public administration* (pp. 510–528). Chicago: Dorsey.

Borins, S. (1995). The new public management is here to stay. *Canadian Public Administration, 38*(1), 122–132.

Osborne, D., & Gaebler, T. (1992). *Reinventing government*. New York: Addison-Wesley.

Peters, B. G., & Savoie, D. J. (1994). Civil service reform: Misdiagnosing the patient. *Public Administration Review, 54*(5), 418.

Rainey, H. G., Backoff, R. W., & Levine, C. H. (1976). Comparing public and private organizations. *Public Administration Review, 36*(2), 233–244.

Savoie, D. J. (1994). *Thatcher, Reagan, Mulroney: In search of a new bureaucracy*. Pittsburgh, PA: University of Pittsburgh Press.

Savoie, D. J. (1995). What is wrong with the new public management? *Canadian Public Administration, 38*(1), 112–121.

Administration Theory

There is no single theory of public administration. Rather, public administration theory contains a complex of statements regarding what administration is, the meaning of *public*, who creates public administration, what the goals of public administration are, and how public administration should fulfill its tasks. Therefore, administration theories have always been and continue to be interdisciplinary, influenced by concepts from history, law, sociology, political science, management, and even psychology. Public administration theory also applies to a distinction between an academic discipline and a profession and contains approaches from organization theory, politics, governance, and law. This entry provides an overview of the development of administrative theories for the public sector and also examines some contemporary challenges in this area. It begins with early theories that sought to distinguish between politics, which dealt with the choice of the values underlying public policy as articulated through the political process, and administration, which implemented those policies in an efficient way similar to the way that businesses operate. It then considers the impact of theories of scientific management on the development of administrative theory. Last, it examines more recent thought that sees the influence of values on administrative decisions and proposes that the distinction between politics and administration is not as clear as it has been considered in the past.

Development of Administrative Theories

The European tradition of public administrative theory was based on Max Weber's conception of bureaucracy as a formal-legal institution responsible for implementing public law. This model was an ideal type that represented the highest form of rational-legal development for the state. Furthermore, it was designed to ensure equal treatment of clients through formal rules and through careful maintenance of files. While bureaucracy is generally a negative term in ordinary usage, its original intentions were more positive. In particular, the bureaucracy concept includes

1. division of labor and division of authorities,
2. hierarchy of administrative bodies, and
3. career system as the basis for the organization.

Public administration in the European tradition has been based on law, with the emphasis on the legal foundations of all public actions. Furthermore, executing public duties defined by law is incorporated into constitutional formula of executive authorities (in countries that have a written constitution), among legislative and judicature. As David Rosenbloom and Robert Kravchuk (2005) note, negatively defining administration emphasizes the autonomy of administrative agencies from legislative and judicature, also indicating that administrative actions affect rights and duties of private subjects. A legal approach to public administration, defined by lawyers, brings out the legal limits of administration and its executive nature; however, as some scientists rightfully state, public administration cannot consist solely of executing legal acts—similarly, it cannot be simple execution of policies, created without its involvement.

The American tradition in public administration has been based more on the work of Woodrow

Wilson and his concern with the relationship between administration and politics. Wilson's famous essay, titled *The Study of Administration* (1887), concentrated on organization and methods of activity of administrative agencies, particularly emphasizing how administrative tasks could be performed to achieve the highest possible efficiency. Wilson also clearly distinguished spheres of politics and administration and initiated the so-called politico-administrative dichotomy. He argued that what distinguishes politics from administration is that administration executes the will of the state as identified through the political process. Numerous scholars such as Frank J. Goodnow, William F. Willoughby, and Leonard White built on Wilson's work to form the beginning of modern administrative theory.

Following Wilson's so-called classic contribution, public administration authorities attempted to develop basic principles of public administration. The classic doctrine assumed that democracy and efficient administration were compatible and that the implementation activities of government could be analyzed scientifically. The early theorists of management have emphasized that organization and control can be recognized almost as synonyms, because almost at the same time the other trend of public administration patterns named scientific management was developing.

Wilson was interested in management science, and Frederick W. Taylor, recognized as the father of management science, began research that resulted in *The Principles of Scientific Management* (1911). The assumption of this work was that there is a single best way to achieve any assigned aim, whereas management science was supposed to help achieve the increase of efficiency by inventing the fastest, the most efficient, and the least "tiresome" production methods. This emphasis on management led to an identification of administrative science with bureaucratic forms of organization. The essential scopes of interests of academics and practitioners in the area of public administration have become division of work, the scope of control, hierarchy of organization, the chain of official internal instructions, the reporting system, departmentalization, development of standards, politics, and activity procedures.

The French engineer Henri Fayol was another significant influence on administrative theory and scientific management. Fayol studied company management seeking to enhance efficiency. According to his theory, administration is recognized as the social science ruled by similar rules to natural sciences. Fayol analyzed administration in terms of the functions of industrial organizations, which he classified as planning, organizing, commanding, coordinating, and controlling. Fayol was a rationalist who understood management of both private sector (companies) and public sector (e.g., public administration) organizations to be exactly the same thing. These rules, which he identified for the private sector, apply to the public one as well.

Classical administrative theory was developed further during the 1930s. This period particularly emphasized the inside aspects of public administration: management methods and problems, structures and activities of organizations, budget procedures, and staff problems. Luther Gulick's approach was contained in an abbreviation—POSDCORB, which management should meet as follows:

Planning—planning, elaborating the direction and methods of planned actions

Organizing—creating formal structure of power, division into organizational units, which cooperate for an assigned aim

Staffing—locating, improving, and developing staff

Directing—ordering, decision-making process, and handling instructions

COordinating—combining separate parts of the same action

Reporting—informing all actors in the process of accomplishing a plan

Budgeting—financial planning, accountancy control

The classical theory was challenged through decision-making approaches, notably those associated with Herbert Simon and bounded rationality. His 1947 book, *Administrative Behavior*, served as the foundation for the development of the so-called Carnegie School that emphasized the limits on rationality and the organizational basis of decisions. This school then served as the foundation for the rebirth of institutional analysis in political science and public administration.

Simon declared that making administrative decisions lies at the heart of public administration, emphasizing the feasibility of using scientific methods to study public administration; he argued that logical positivism was the right approach in terms of creating administrative policies. He sought to develop a social scientific approach to administration rather than the principles associated with the then conventional wisdom of "principles of administration." Simon disapproved Gulick's approach and his POSDCORB, claiming it to be inappropriate for many situations that civil servants face. He created the idea of limited rationality in the decision making, recognizing that people are rational in making decisions, but only to a certain extent. Simon acknowledged that decisions made by civil servants are rational in their view, but at the same time he pointed to many alternative variants that lead to achieving the most rational decision. According to Simon, the new paradigm for the study of public administration means that there should exist two trends in the discipline, harmonizing and mutually, intellectually stimulating:

1. One trend concentrated on progress of pure administrative science, strongly based on fundamentals of social psychology to understand decision making more effectively.
2. Another, more representative trend, aimed at creating guidelines of public policies that could guide decisions, albeit without the understanding of decision-making theory implied in the first principle.

Robert A. Dahl (1947), one of the most important early behaviorists of political science, analyzed "state of the art" of administrative science. According to him, administrative science should

1. recognize the complexity of human behavior,
2. relate to the problems of values eligible in administrative situations, and
3. consider the connections between public administration and surrounding environment.

Critics of Simon's fundamentals of public administration and political-administrative dichotomy were also supported by another scientist, Dwight Waldo, who defined public administration from the perspective of the surrounding environment, emphasizing the importance of comparative administration. From Waldo's point of view, public administration should be a branch of political science.

The more political approach to public administration has led to the conclusion that public administration is not only accountable to elected public officials (members of parliament and executive authorities) but also responds to their needs, in other words, being an instrument in politicians' hands. The importance of public administration from this perspective is stressed by the fact that it has the discretional authority. B. Guy Peters considers public administration to have a broader scope than a typical bureaucracy (in terms of random organization of administrative character) and emphasizes that it is an integral part of decisive processes of executive authorities.

Contemporary Challenges

Public administration today tends to equate public administration and governance. Etymologically, the term *governance* derives from the Greek word *kubernân* (piloting, steering) and was used by Plato to describe the creation of systems of rules. The Greek source gave birth to the Latin *gubernare*, with a similar meaning (piloting, making rules, steering). Governing, as we make a reference to it here, is close to the term of *managing* (leading), conducted by governments and their administration. Governing in this approach is connected to the new forms of actions of public administration, which are difficult to classify strictly as executing the law. For example, the World Bank created a definition of "governance" meaning the traditions and institutions by which authority is exercised for the common good. This approach, portraying "different" governing (with regard to traditional administration), is close to the new public management (NPM) conception and forms the base for further analysis of public administration and making new trends, such as new public service (NPS).

The NPM has been a recent and pervasive counterpoise to the emphasis on law in most continental (and Latin American) administrative systems. This approach to public administration emphasizes the economic foundations of governing and argues that public and private management are essentially the same. Therefore, this theory of administration in the public sector separates politics and administration even more than in the traditional Wilsonian and Weberian formulations. The NPM approach represents a recent theory for public administration, but as its influence wanes at the beginning of the 21st century, there is as yet no clear replacement. That said, the political and managerial weaknesses inherent in NPM have produced a number of important reactions in practice, if not in theory.

Patrycja J. Suwaj
University of Bialystok
Bialystok, Poland

See also Bureaucracy; New Public Management; Rationality, Bounded

Further Readings

Christensen, T., & Laegreid, P. (2007). *Transcending new public management.* Aldershot, UK: Ashgate.
Dahl, R. A. (1947). The science of public administration: Three problems. *Public Administration Review, 7,* 1–11.
Peters, B. G., & Pierre, J. (2005). *Handbook of public administration.* London: Sage.
Rosenbloom, D. H., & Kravchuk, R. S. (2005). *Public administration: Understanding management, politics and law in the public sector*. New York: McGraw-Hill.
Simon, H. A. (1947). *Administrative behavior*. New York: Free Press.
Waldo, D. (1955). *The study of public administration*. Garden City, NY: Doubleday.

ADVOCACY

Advocacy is a kind of political action addressed to a governing body with the aim of influencing public policy outputs. Though advocacy is often equated with lobbying, it is in fact an umbrella term for organized activism related to a particular set of issues. Lobbying, meanwhile, is the practice of influencing a governing body through direct communication with legislators. Advocacy groups strive to change public policy according to their interests without seeking political mandates. The definition emphasizes the existence of some kind of organization, ranging from loosely coupled networks to highly formalized nonmembership organizations. It includes social movements, voluntary associations, or membership groups as well as institutions such as corporations, universities, cities, or foundations. This entry discusses the major actors, targets, and strategies, and their respective impacts.

Actors

The increase in the number of advocacy groups is one of the most striking and still ongoing trends in Western societies. The number of citizens willing to join an interest group has increased over many decades. In the United States, which has been in the vanguard, this so-called advocacy explosion started around 1960. Whereas between 1920 and 1960 the number of interest groups represented in Washington merely doubled, it quadrupled within the following 20 years. Similarly, at the end of the 1980s, the number of interest groups represented at the European Union (EU) in Brussels was 10 times higher than it was at the beginning of the 1970s, and it quadrupled between 1985 and 1991. The rapid growth in number was accompanied by a structural change with respect to the scope, intensity, and shape of advocacy. Today, advocacy is more diverse and much more professional than it was a few decades ago, including an increase in staff acting as in-house lobbyists, a growth of different kinds of lobbying firms performing surrogate representation, and the development of more and more forms of direct and indirect lobbying.

In the first half of the 20th century, the universe of advocacy groups in the Western world was dominated by trade, professional, and labor organizations. For example, before 1950 in the United States, these three types of groups accounted for over three quarters of all organized interests active in politics. Beginning in the 1960s, however, we can find a significant decline in the rate of unionization, first of all in the United States, and, since the 1980s, also in most Western European countries. Although the number of the other types of established groups continued to increase even in the second half of the century, their share of the group universe declined due to an even more rapid growth of various types of organizations from the nonprofit sector concerned with health, social welfare, culture, education, public affairs, consumers, senior citizens, taxpayers, civil rights, women's rights, human rights, the environment, religion, and other areas. Many of these groups originated in the social movements of the 1960s and 1970s, which mobilized their constituencies to turn the public's attention to new interests and new political problems. Parts of these movements turned into associations that would allow the groups to maintain their capacity to advocate even in times of declining mobilization.

The change in the composition of the group universe was paralleled by a process of differentiation in the mode of organization. The advocacy explosion resulted not only from a rise in the number of voluntary associations but also—and perhaps even more—from the fact that more and more institutions, such as corporations, governmental entities, universities, hospitals, and think tanks, have begun to act as interest groups. Many advocacy groups thus no longer represent the aggregate interests of individual group members but are lobbying on behalf of corporate actors, which do not have any members at all. Today, we find a highly differentiated pattern of advocacy in almost all Western countries consisting of a great variety of types of organizations, including broad-based membership groups such as labor unions, professional groups, and citizen groups, as well as institutions and groups with corporate members, such as trade associations and complex umbrella associations.

The different developments in the group universe were induced by several factors. First of all, the rapid social and economic changes in Western countries have created rising numbers of potential members and an overall growth in organizational resources. As people have become more educated, more articulate, and wealthier, and as knowledge and information have become more widespread, new interests have emerged, and traditional ones have been redefined. Accelerated by advanced communication technologies, the translation of interests into formal group organizations became a cumulative process, stimulating other constituencies to organize or to amplify political activities. Recently, another determinant of the expansion of the advocacy group system has come to the fore: government activity. Whenever government activity is expanded to new issues or existing government bodies gain new jurisdictions, new groups will form or existing ones will focus their activities on these new political areas. In spite of these driving forces, advocacy growth may slow down or reach a limit in the future, as it will become increasingly difficult for organizers to find interests not yet politically represented by any group or to mobilize for advocacy in view of scarce resources.

Levels and Targets

The targets of advocacy differ depending on the institutional settings and the decision-making procedures of the political system. Specific structures such as the separation of powers or the formal and informal rules of policy making are critical to advocacy groups' decisions as to whom to lobby and where to try to gain access. Whereas in centralized systems advocacy will be focused on actors and institutions of the central authority, multilevel or federated systems show a much more differentiated pattern of advocacy. For example, despite Germany being a federal state, advocacy groups in Germany focus on the federal level, particularly on the ministries and on leading parliamentarians, such as the chairs of committees, due to the strong and centralized party system. In contrast, the American federal system, even at the local level with its more than 80,000 governments, forms a significant political target since legislators are heavily dependent on the financial and political support of local coalitions. Finally, in the highly differentiated multilevel system of the EU, advocacy groups find many points of access and several veto points, offering a variety of opportunities to exert influence on policy outcomes.

The choice of political targets also depends on the character and type of the advocacy group. While traditional interest groups are apt to focus their activities on the national level, nongovernmental organizations (NGOs), most of them representing the interests of disadvantaged groups, focus on the international level, where their concerns have a better chance of being accepted than at the national level. Thus, Eurogroups tend to be experts in multilevel strategies, acting at the European as well as on the respective national levels, depending on which of the two is most critical for the decision-making process in which they engage. Legislators and governing boards, however, not only are passive targets of advocacy groups, but they often have an impact on group structures and resources as well. The most common instrument is the financial support given to groups, which otherwise would not be able to organize. Such support is intended to create countervailing forces to the powerful interest groups that tend to dominate the political process or to help form coalitions of advocacy groups that will improve the legislators' chances of getting reelected or foster their political plans.

Social change and changes in technology, institutional settings, and decision-making procedures will often prompt advocacy groups to address new targets and different levels of the political system. As the European Parliament has gained increasing power and responsibilities, it has increasingly become a target of lobbying activity. Similarly, globalization and denationalization have created new opportunity structures for social movement organizations and NGOs. Relying on new communication technologies, they have developed transnational ties and coordination mechanisms that enable them to organize protest on the international level and to develop countervailing power to the policies of international institutions such as the World Trade Organization or the International Monetary Fund. But these processes of adaptation may fail if the characteristics of the advocacy groups do not correspond to the institutional context in which a group wishes to act. For instance, the lack of a genuine European public sphere and the specific decision-making style of the European institutions involve high transaction costs for the mobilization of mass protest in Brussels. Therefore, the activities of social movements have remained predominantly national up to now.

Strategies and Tactics

To begin with, advocacy includes many activities that seem to have little to do with lobbying at all, such as gathering information, monitoring the political process, and conducting research. Nevertheless, today we find a great variety of tactics and strategies applied by advocacy groups and lobbyists to make decision makers willing to produce policy outputs consistent with their interests. The single group, however, has to decide which tools to apply and which not, since resources are limited. As a rule, advocacy is focused on a small number of subjects. Lobbyists spend most of their time on only a few critical issues, whereas they spend little time on many others.

For a long time, scholars have agreed that one should distinguish between inside strategies (based on personal contacts between lobbyists and policymakers) and outside strategies (public relations and grassroots contacts). This distinction has been useful for finding correlations between group characteristics or other factors and the selection of tools and for describing shifts in the strategic orientations of advocacy groups. There is a lot of evidence demonstrating that citizen groups will prefer outside strategies while established groups, mainly business, rely heavily on direct lobbying. Recently, however, the overall picture has become more multifaceted. While business groups tend to apply multivoice strategies that give them the chance to be present wherever a channel of influence opens up, social movements and citizen groups have also tried to apply inside strategies. Considering the growing variety of strategies and tactics, the distinction between inside and outside strategies seems to become obsolete and should be replaced by a more differentiated typology as suggested by Jeffrey Berry, who identifies four different advocacy group strategies, namely,

1. litigation and administrative intervention,
2. confrontation (protest, whistle-blowing, public relations, etc.),
3. information (making personal representations to government, releasing research results), and
4. constituency influence.

One of the major factors that has accounted for an expansion of lobbying activities is advocacy explosion. The rise of countervailing power has strengthened competition between new and established groups, forcing the latter to spend more on lobbying and to expand their action repertory. Furthermore, technological innovation in information processing has created new opportunities for lobbyists to gain access to policymakers and to apply new forms of lobbying, particularly those that address the general public, such as media- or issue-based campaigns. These trends have made it necessary for advocacy groups to acquire new skills and more professional knowledge, required to apply the new tools in an efficient and effective way. The rising demand for advocacy advice, in turn, is the main reason why direct representation by in-house lobbyists has increasingly been supplemented by surrogate representation through lobbying firms. Acting in the name of advocacy groups, such firms can help manage the problems posed by new developments in politics and communication technology. Lobbying firms are common primarily in the United States, where they account for about 20% of all lobbyists present in the capital and where they form a special sector with different types of businesses. In the past 2 decades, we have also seen a rapid growth of lobbying firms in the EU. The increasing scope and density of EU regulations have created a rapidly rising demand for professional advice on how to effectively lobby the European institutions.

Structures

Advocacy not only involves unilateral political activities by single groups or their agents but usually results in more or less intensive and stable relationships with other groups and with the groups' counterparts in the legislative and executive branches of government. These relationships come about if advocacy groups have at their disposal certain resources—be it power based on the control of their membership or information that their counterparts are interested in—and, in turn, if other groups or policymakers can offer them valuable information or give them the opportunity to get directly involved in the formulation and implementation of policies. One basic element of the structure of advocacy is thus an exchange relation between private and state actors. In a given policy sector, all the actors that are connected in this way form a specific policy network characterized by the number and type of its members, its stability, and its structure. Basic elements of a political system, such as the characteristics of government and the mode of interest group participation in decision making, form a framework that promotes the development of certain types of networks and helps them persist. By contrast, as competition between groups has increased, the established actors have been forced to open up policy networks and to integrate new advocacy groups in the policy-making process.

Since decision making in contemporary politics is a very complex process, the scope and density of policy networks vary to a considerable extent, making it an empirical question how stable the relationships between actors are and how power resources are distributed. Most scholars distinguish between several types of networks that can be arranged in a continuum ranging from very open to rather closed patterns of interaction. Whereas in the American context scholars tend to distinguish between "iron triangles" or subgovernments at the one extreme and issue networks at the other, the European debate on policy networks is shaped by the opposition of corporatism and pluralism. In spite of these differences in terminology, descriptions of the two extreme types are similar. While the former is characterized by highly restricted participation, stability of interactions, and centralized decision making, the main features of the latter are decentralization, a high level of conflict, dynamism, and broad participation.

The academic discourse on pluralism and corporatism is characterized by a typical asymmetry. As Grant Jordan stated, pluralism is ultimately no more than an antitheory, which is marked by what it excludes rather than by what it establishes. First of all, it describes a complex and confusing configuration of political actors that is open for new advocacy groups and where power is widely dispersed. Its core element is open competition between advocacy groups for political influence, which prevents any single group from dominating the political process. Since pluralism involves a great variety of single acts of exchange between private groups and legislators, the results of political decision making are highly unpredictable. In the corporatist pattern, however, a small number of advocacy groups have considerable resources at their disposal, in particular technical information on the implementation process that makes them essential partners of the legislators, with whom they work in close coordination. Corporatist arrangements consist of stable sets of actors in which the privileged advocacy groups get the chance to considerably influence policy outcomes in return for their cooperation in the implementation of a policy. As less resourceful societal actors are excluded, corporatism tends to subdivide the group universe into insiders and outsiders.

Whether corporatist arrangements have a chance to develop depends on the institutional, political, and social context. In political systems with a strong central state and encompassing umbrella organizations, as in several European countries, policy processes have been dominated for decades by interelite relations often based, as in Germany, on the public status of the major private actors. By contrast, in the United States, the fragmentation of economic interest groups and the division of American government has prevented corporatist patterns from developing. Since labor unions have always been weak, there was no strong stimulus for business organizations to develop powerful peak associations. So competition between groups of labor and capital has remained high and discouraged single groups from becoming privileged partners of the legislators. Besides, the American political system offers little opportunity to channel advocacy group activities in a way that would favor the development of corporatist arrangements. Federalism, the separation of powers, the fragmentation of power within both Congress and the executive, and the lack of a strong system of party discipline are factors that favor pluralist rather than corporatist patterns of interest intermediation.

Impact

In the modern world of politics, the political activity of advocacy groups is only one of a multiplicity of factors having an impact on public policy decisions. Therefore, methodological approaches that attempt to isolate this factor from the institutional and situational context in which it is embedded have produced unsatisfactory results. Research on advocacy group influence came up with better results when it stopped asking whether or to what extent advocacy groups are powerful at all and instead turned to the less fundamental question of under which circumstances they have an impact on policy outcomes. Today we are able to identify the structural and situational factors that help explain why advocacy is successful in some cases but in others has no impact at all. The five determinants are as follows:

1. group-related characteristics such as size and type of membership; resource endowment, particularly with regard to money; and the tactical and strategic repertory;
2. the features of the group universe, that is, the number of groups in a policy sector and the degree of cooperation and conflict;
3. legislators' demand for group resources such as information and political support;
4. the institutional framework as a channel for the distribution of opportunities to gain access to decision makers or to gain legal privileges; and
5. issue characteristics, such as policy type, degree of technicality, and public salience.

At first, it seems reasonable to assume that, other factors being equal, advocacy groups that are well endowed with different resources will be more influential than groups that are less endowed. The former will be able to employ more staff and pursue more multifaceted tactics and strategies than the latter. Furthermore, it is often suggested that an imbalance of influence is due to the unequal distribution of money. Indeed, there are many examples of rich groups, particularly corporations and trade associations, gaining easier access to decision makers, being present at more levels of government, and supplying decision makers with more valuable expertise than poorer groups are able to do. But there are only few academic studies that would confirm the assumption that political outcomes are skewed toward the interests of the advocacy groups that spend the most on lobbying. Money has only a limited effect, since there are other resources, such as public support, a large number of members, or political credibility, that may outweigh material affluence and, still more important, that cannot be bought on the political market. Moreover, recent research results suggest that group-related factors, such as money and membership and also tactics and strategies, all in all may have only a limited impact on political outcomes.

Keeping the other factors in mind, these findings do not come as a surprise. First of all, in a competitive group universe most attempts by advocacy groups to exert influence evoke counteracting activities by rival groups. For example, a sector such as business is, in most cases, unable to act as a unit for it is divided by too many conflicting interests. Furthermore, the advocacy explosion has intensified competition by creating more and more countervailing pressures. Second, advocacy success is determined by the behavior of legislators. Whereas legislators are often considered as passive targets, in fact they are active players shaping their relationship with advocacy groups according to their interests. Whether a decision maker is influenced by a certain group depends on his or her political goals, strategies, and demand for resources. Third, in any political system, institutions form a screen that is beneficial to some advocacy groups and discriminates against others. Generally, institutional mechanisms may preclude advocacy groups from converting resources into power. If the structures and strategies of a group do not match up with the political institutions, as is the case with social movements on the EU level, even resourceful groups may remain politically ineffective. Fourth, the likelihood of an advocate prevailing in a political conflict also depends on the issue at stake, that is, on the constellation of actors interested in the issue and their positions on the subject. The more groups that are involved in the issue and the higher the level of conflict, the more interests legislators have to take into account. So big issues have an adverse effect on lobbying success.

As a result, the impact of advocacy is limited, depending on factors such as the scope of the policy network, the level of conflict, the number of allies and opponents, and so on. Since every policy process is characterized by a specific, if not unique, constellation of structural and situational elements, the effect of group tactics and strategies is highly uncertain. Provided that the combination of factors in a certain policy arena is constantly changing, as is often the case, the results of public policy decisions are highly unpredictable, making it extremely unlikely for single groups to dominate the political process over time. But we will also find political contexts with a smaller number of actors and more stable patterns of participation where policy outcomes are less indeterminate. Advocacy groups that are formally included in policy-making processes, that deal with an issue that does not attract the interest of the media, and that act in a political niche where countervailing pressures are low will have a much better chance to influence policy outcomes than groups that act under largely varying conditions. As structural and situational factors can be considered as variables, the level of uncertainty is a variable too, featuring different contexts that are sometimes favorable to some groups and sometimes to others.

Thomas von Winter
University of Potsdam
Potsdam, Germany

See also Advocacy Coalition Framework; Interest Groups; Lobbying; Neo-Corporatism; Pluralist Interest Intermediation

Further Readings

Baumgartner, F. R., & Leech, B. (1998). *Basic interests: The importance of groups in politics and political science*. Princeton, NJ: Princeton University Press.

Berry, J. M., & Wilcox, C. (2007). *The interest group society* (4th ed.). New York: Pearson Longman.

Eising, R., & Kohler-Koch, B. (Eds.). (2005). *Interessenpolitik in Europa* [Interest politics in Europe]. Baden-Baden, Germany: Nomos.

Heinz, J. P., Laumann, E. O., Nelson, R. L., & Salisbury, R. H. (1993). *The hollow core: Private interests in national policy making*. Cambridge, MA: Harvard University Press.

Herrnson, P. S., Shaiko, R. G., & Wilcox, C. (Eds.). (1998). *The interest group connection: Electioneering, lobbying, and policymaking in Washington*. Chatham, NJ: Chatham House.

Lahusen, C., & Jauß, C. (2001). *Lobbying als Beruf: Interessengruppen in der Europäischen Union* [Lobbying as a profession: Interest groups in the European Union]. Baden-Baden, Germany: Nomos.

Mahoney, C. (2007). Lobbying success in the United States and the European Union. *Journal of Public Policy, 1*, 35–56.

Nownes, A. J. (2001). *Pressure and power: Organized interests in American politics*. Boston: Houghton Mifflin.

Warntjen, A., & Wonka, A. (Eds.). (2004). *Governance in Europe: The role of interest groups*. Baden-Baden, Germany: Nomos.

Advocacy Coalition Framework

The advocacy coalition framework (ACF) was originally designed by Paul Sabatier and Hank Jenkins-Smith in the late 1980s to explain the political behavior of actors in the policy process.

The ACF provides researchers a theoretical lens for organizing actors into coalitions by offering a set of hypotheses for explaining coalition structure and behavior, the role of science and technology in the policy process, and the factors conducive to learning and policy change. For more than 2 decades, the ACF has been applied in Europe, Canada, the United States, and developing countries to a range of topics including health policy, environmental/energy policy, economic policy, disaster policy, and education policy.

Sabatier and Jenkins-Smith created the ACF in response to what they saw as three shortcomings in the policy process literature. The first was their interpretation of the stages heuristic as both an inadequate causal theory and an inaccurate portrayal of the policy process. The second was a lack of progress to resolve the debate about top-down and bottom-up approaches to implementation research and a need for system-based theories of policy making. The third was the apparent lack of theory about the role of scientific and technical information in the policy process. In response, the ACF was created as a system-based model, integrating most of the policy stages, incorporating the best aspects of the top-down and bottom-up approaches to implementation, and creating hypotheses that involve the role of scientific and technical information in learning and policy change. This entry provides a broad overview of the assumptions, concepts, and hypotheses in the ACF and ends with directions for future research.

Foundation

The structural underpinnings that the ACF begins with are a set of assumptions involving a central role for scientific and technical information in policy processes, a time perspective of 10 years or more to understand policy change, and the assignment of policy subsystems as the primary unit of analysis. Policy subsystems are subunits of political systems, focused on a topical area and involving specialized actors. The ACF guides researchers to study a broad set of subsystem actors including officials from all levels of government, nongovernment actors, consultants, scientists, and members of the media. For these subsystem actors, the ACF assumes a mental model based on limited abilities to process stimuli, a tendency to remember losses more than gains, and a need to rely on beliefs as the principal heuristic to simplify, filter, and sometimes distort stimuli.

The ACF assumes that actors are politically driven by their beliefs and that their policies and programs are best thought of as translations of those beliefs. A three-tiered model of a belief system for its actors is depicted. Within the belief-systems model, the broadest, most rigid, and predominately normative are *deep core beliefs*, such as liberal and conservative beliefs or the relative concern for the welfare of present versus future generations. Of moderate scope and spanning the substantive and geographic breadth of a policy subsystem are *policy core beliefs*. Policy core beliefs are rigid but are more adaptive in response to new experiences and information than deep core beliefs. The most substantively and geographically narrow and the most empirical are *secondary beliefs*. Compared with deep core and policy core beliefs, secondary beliefs are more likely to change.

At the heart of the ACF is the coalition concept. Coalitions consist of actors with shared policy core beliefs who engage in nontrivial degree of coordination, ranging from developing joint plans to sharing information. The goal of a coalition is the attainment of policy objectives, meaning the protection or change in government policies. The behavior of a coalition toward the achievement of their policy objectives varies considerably and depends on their resources and on the resources of their opponents, all of which is constrained by external subsystem affairs. Common coalition strategies include (a) maintaining a long-term perspective and involvement in a policy subsystem, (b) seeking to maintain membership and mobilizing new members, (c) attempting to shape policies indirectly through public opinion campaigns and through supporting scientific and technical information that reinforces their beliefs, and (d) attempting to shape policies directly by shopping their ideas to venues controlled by actors with authority to change policies, such as the courts, executives, agencies, and legislatures. The behavior of coalitions is fundamentally connected to the level of conflict in the policy subsystem. In some subsystems, rivals are not actively present, and a dominant coalition maintains a monopoly over policy affairs. In more adversarial subsystems, coalitions behave caustically toward their rivals, limiting learning between coalitions. In more collaborative subsystems, coalitions cooperate, share information, and engage in cross-coalition learning.

While coalition members share policy core beliefs and coordinate their behavior, not all coalition members need to be directly tied to their allies. Some members are principal members, who remain active within the coalition for extended periods of time and are well connected with most allies. Principal members anchor the coalition by bearing the cost of coordination and by maintaining connections with—and hence the involvement of—more auxiliary members, whose involvement tends to be more temporally sporadic.

A flow diagram of the ACF is shown in Figure 1. On the right is a policy subsystem, within which coalitions strive to attain their policy objectives through various strategies. Subsystems operate within a broader political environment defined by relatively stable parameters and external events along with long-term coalition opportunity structures and short-term constraints and resources of subsystem actors.

To ease the application across political systems, the flow diagram in Figure 1 depicts three sets of variables as important long-term opportunity structures. The first is the degree of consensus needed for major policy change, which shapes coalition strategies in reaching agreements. The second is the degree of openness of political systems. For example, federalism and checks and balances in the United States create decentralized processes with many venues and encourage entry participation whereas corporatist systems are less open and more centralized and restrict participation. The third is overlapping societal cleavages.

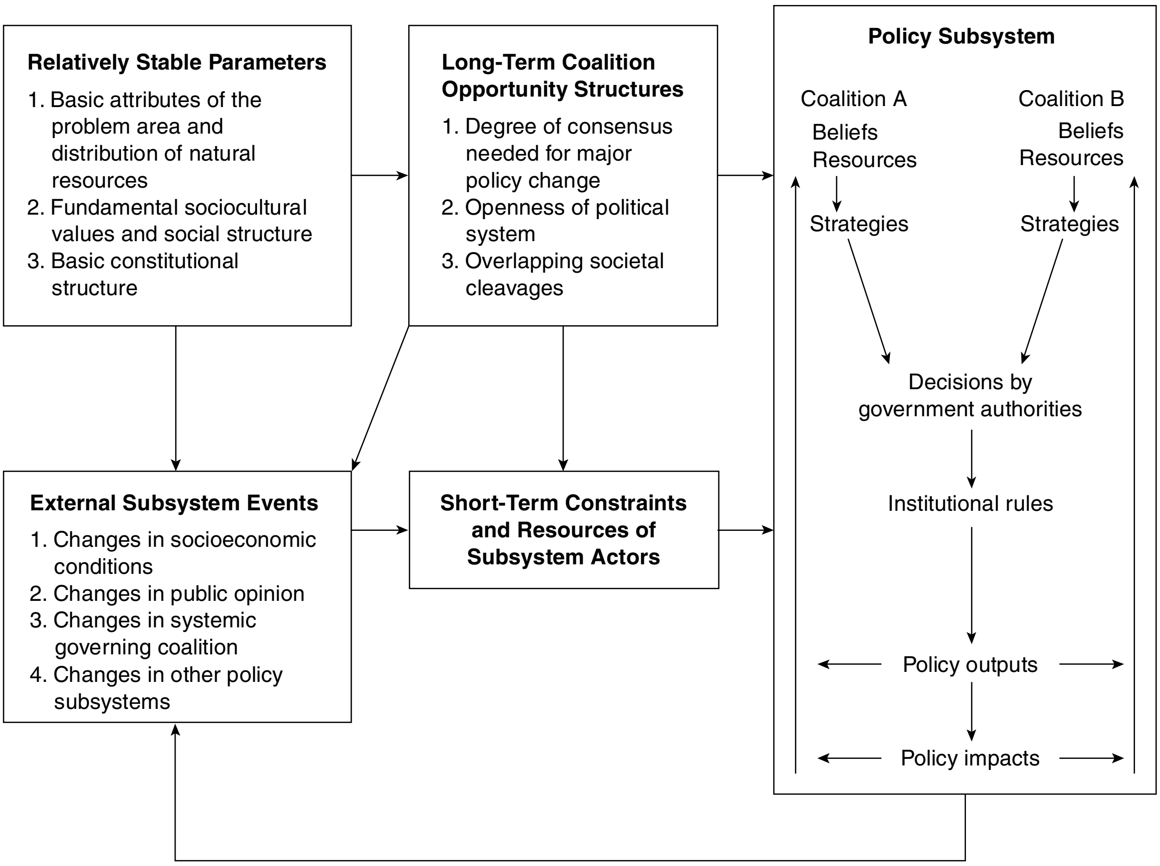


Figure 1 A Flow Diagram for the Advocacy Coalition Framework

Source: Adapted from Sabatier, P. A., & Weible, C. M. (2007). The advocacy coalition framework: Innovations and clarifications. In P. A. Sabatier (Ed.), *Theories of the policy process* (2nd ed., pp. 189–222). Boulder, CO: Westview Press.

Paths to Major Policy Change

An explanatory focus of the ACF is to help understand and explain major policy change. Four paths to major policy change have been posited. The first is shocks that occur external to a subsystem, including broad changes in socioeconomic conditions, public opinion, governing coalitions, and other subsystems. Among the factors linking an external shock to major policy change are shifting coalition resources and changing beliefs among coalition actors.

The second path to policy change is shocks within a policy subsystem. An example might include a catastrophic fire in a forest management policy subsystem. The ACF distinguishes between external shocks and internal shocks because the latter are predicted to underscore the failures in current subsystem practices and possibly weakening a dominant coalition.

The third path to policy change is policy-oriented learning. Policy-oriented learning involves relatively enduring changes in beliefs and strategies resulting from experience and/or new information as related to the attainment, and possibly the modifications, of policy objectives. As coalition members learn over time, they may develop new ways to overcome a rival coalition or refine their policy objectives, both of which may lead to major changes in subsystem policies. Since belief systems are rigid and resistant to change, policy-oriented learning primarily shapes secondary beliefs or secondary aspects of the policy subsystem over extended periods of time.

The fourth path to policy change occurs through negotiated agreements among competing coalitions. Probably, the most important condition enabling coalitions to negotiate is a hurting stalemate, which occurs when members of rival coalitions view the status quo as unacceptable and perceive no other venues to achieve their policy objectives other than direct negotiations. Additional conditions enabling negotiated agreements include effective leadership, consensus-based decision rules, diverse funding, duration of process and commitment of members, a focus on empirical issues, an emphasis on building trust, and lack of alternative venues.

Hypotheses

Sabatier and Jenkins-Smith provide a dozen hypotheses on topics ranging from belief and policy change to coalition stability. The five most frequently tested hypotheses are the following:

1. Significant perturbations external to the subsystem (e.g., changes in socioeconomic conditions, public opinion, systemwide governing coalitions, or policy outputs from other subsystems) are a *necessary*, *but not sufficient*, cause of change in the *policy core* attributes of a governmental program.
2. On major controversies within a policy subsystem when policy core beliefs are in dispute, the lineup of allies and opponents tends to be rather stable over periods of a decade or so.
3. Policy-oriented learning across belief systems is most likely when there is an intermediate level of informed conflict between the two coalitions. This requires that (a) each has the technical resources to engage in such a debate and (b) the conflict be between secondary aspects of one belief system and core elements of the other or, alternatively, between important secondary aspects of the two belief systems.
4. Policy-oriented learning across belief systems is most likely when there exists a forum that is (a) prestigious enough to force professionals from different coalitions to participate and (b) dominated by professional norms.
5. The policy core attributes of a governmental program in a specific jurisdiction will not be significantly revised as long as the subsystem advocacy coalition that instituted the program remains in power within that jurisdiction—except when the change is imposed by a hierarchically superior jurisdiction.

Further Development and Applications

The scope of the ACF continues to diversify in topics and theoretical focus. Areas ripe for future research include (a) explaining coalition structure and behavior, especially factors explaining defection and stability of members and choice of political strategy; (b) developing a theory of policy subsystem interdependencies; (c) explaining the effects of scientific and technical information on learning within and between coalitions; and

(d) investigating how biased assimilation shapes coalition behavior.

Christopher M. Weible
University of Colorado Denver
Denver, Colorado, United States

Paul A. Sabatier
University of California, Davis
Davis, California, United States

See also Agenda Setting; Environmental Policy; Network Analysis; Policy Analysis; Policy Cycle; Policy Process, Models of; Social Movements; Values

Further Readings

Kübler, D. (2001). Understanding policy change with the advocacy coalition framework: An application to Swiss drug policy. *Journal of European Public Policy, 8*(4), 623–641.

Sabatier, P. A. (1986). Top-down and bottom-up approaches to implementation research: A critical analyses and suggested synthesis. *Journal of Public Policy, 6,* 21–48.

Sabatier, P. A., & Jenkins-Smith, H. (1993). *Policy change and learning: An advocacy coalition approach.* Boulder, CO: Westview Press.

Sabatier, P. A., & Jenkins-Smith, H. (1999). The advocacy coalition framework: An assessment. In P. Sabatier (Ed.), *Theories of the policy process* (pp. 211–236). Boulder, CO: Westview Press.

Sabatier, P. A., & Weible, C. M. (2007). The advocacy coalition framework: Innovations and clarifications. In P. Sabatier (Ed.), *Theories of the policy process* (2nd ed., pp. 189–222). Boulder, CO: Westview Press.

Weible, C. M. (2007). An advocacy coalition framework approach to stakeholder analysis: Understanding the political context of California marine protected area policy. *Journal of Public Administration Research and Theory, 17*(1), 95–117.

Advocacy Networks, Transnational

Transnational advocacy networks (TANs) are forms of collective action that cross national borders and are made up of collaborative connections among a myriad of actors who engage in joint initiatives on the basis of common goals and shared values. These networks seek to change states' and/or international organizations' policies and/or public opinion, at both a domestic and an international scale. TANs form around areas of specific issues, but these vary considerably, ranging from human rights, environmental protection, poverty, gender equality, and labor rights to networks that seek to change the agendas and policies of international organizations such as the World Bank or the International Monetary Fund. Their initiatives usually have a limited life span. Its promoters do not seek to build long-term coalitions; rather, TANs emerge and die as new issues come to the fore or as their demands gain attention.

The concept of TANs thus describes a different form of organized collective action from the more institutionalized international nongovernmental organizations or global social movements. Although some kind of internal division of labor and communication structure is in place, TANs most often do not have headquarters or paid staff. Furthermore, rules for affiliation tend to be loose. However, the internal functioning of TANs varies considerably, and it changes through time and across issue areas.

The literature on TANs is quite recent. It flourished in the 1990s, as part of a broader effort to understand the roles of nonstate actors in international relations. Ever since the publication of the pioneering analyses on transnational relations, around the end of the 1960s and the beginning of the 1970s, its study has gained increased prominence. What began as an attempt to incorporate the study of nonstate actors—at the time, mainly multinational corporations—into the research agenda of international relations scholars has become a vast literature that takes into consideration an increasingly heterogeneous set of actors and phenomena. The literature on TANs has contributed to moving the debate from the question of whether nonstate actors are relevant in international relations to understanding how they are relevant and how to evaluate their impacts.

Studies about TANs are often multidisciplinary. More specifically, this is a literature that has stimulated a fruitful dialogue between international relations and social movement theories. Although most authors will deny the existence of a simple causal link between globalization and transnational collective action, the constitution of international regimes and the augmented power of international organizations provide a common political background for TANs. These are understood best by considering the contributions made by international relations theorists. In turn, many of the challenges that social movements face at the domestic level, in terms of coalition building, for example, are also present in the international arena. Most important, social movement theorists help in identifying the mechanisms by which actors are able (or unable) to overcome their differences and construct common purpose.

Studies about TANs also seek to go beyond local–global dichotomies. Thus, rather than speak of the emergence or existence of a "global civil society," scholars refer to a plurality of TANs springing up intermittently across a wide-ranging set of issues. These networks are made up of various combinations in terms of participants' national origins and the pathways to transnationality chosen by them. While many actors remain rooted at the local or national scale, others may be part of global social movements or international nongovernmental organizations.

At least two issues remain as key challenges for the TAN literature. First, it is not always clear as to how to set the boundaries of TANs, that is, how to identify participants and differentiate among actors, given the internal asymmetry of power relations. Second, although there has been a concerted effort to evaluate the impacts of these collective action forms, we still know little about their longer term effects.

Relationships Among Actors in TANs

TANs are sponsored by a wide variety of actors located in different social positions and geographical locations, so much so that it is often hard to draw their boundaries. There is no consensus in the literature on this topic. For some scholars, TANs are not necessarily limited to nonstate agents but may include sympathetic government officials as well as international organization officials. Among nonstate actors, these may include nongovernmental organizations (NGOs), grassroots organizations, foundations, business organizations, the media, churches, and various other types of associations as well as individuals. For other authors, it is best to define TANs, more restrictively, as including only domestic and international nonstate actors, and among these mostly NGOs and social movements. A clearer definition of the boundaries of TANs remains an open challenge.

Whether the definition of participants is more or less restricted, in both cases, there is evidence that actors have very different approaches in terms of targets chosen and strategies used. As sets of interconnected nodes, networks cannot be aprioristically defined as formed by horizontal or homogeneous ties. Nevertheless, this use of the concept of "networks" has been common among activists and scholars alike. It has helped emphasize the novelty of these forms of collaboration with respect to more hierarchical and centralized initiatives of the past. However, recent scholarly efforts have focused on trying to specify better the embeddedness of networks in power relations as spaces of negotiation (and contention) affected by asymmetrical relations among participants. These efforts have led to a more detailed analysis about the internal functioning of TANs, which focus, for example, on questions of accountability, transparency, and decision-making processes. Asymmetries may follow North–South lines, but this is only one possible aspect. Other sources of tension within TANs refer to different political cultures (among countries and also among sectoral domains) and different visions in terms of tactics, both of which can be found within and across the South and the North.

Impacts of TANs

To understand under what domestic and international circumstances do TANs succeed or fail to achieve their goals, a long series of case studies and comparisons among them have been undertaken. Scholars have focused on the abilities of networks not only to influence political actors and policies at the domestic and international scales but also to bring about shifts in public opinion on given issues. Research has been done on campaigns oriented toward particular political problems (e.g., debt relief, use of land mines) and policy domains (e.g., the environment, trade negotiations, gender equality). While there are numerous case studies showing that TANs did have a positive impact, this is often not the case. Independent variables that explain these different outcomes in the literature include characteristics of the networks themselves (e.g., the strength of local grassroots members or the density of ties), characteristics of the issue chosen (e.g., its resonance with existing norms), characteristics of the target (e.g., the vulnerability of states to leverage politics or the access of activists to the state), and the strategies of the networks (e.g., their capacity to generate reliable information quickly).

Margaret Keck and Kathryn Sikkink pioneered the efforts to understand the impacts of TANs in their 1998 book, *Activists Beyond Borders: Advocacy Networks in International Politics*, in which they propose the "boomerang pattern" as a common and effective tactic used by TANs to enhance their influence. Most especially in cases of closure of the domestic political system, nonstate actors seek allies in other countries (or are sought by them) to apply pressure on other states and/or international organizations. These, in turn, will be able to use their leverage on the original target and thus help bring about the change demanded by domestic activists. Examples of the successful throwing of "boomerangs" include the well-known case of the Brazilian rubber tappers who, in the 1980s, allied with U.S. environmental NGOs to pressure the U.S. Congress to use its leverage on the World Bank. The World Bank, in turn, helped bring about changes in the projects they helped sponsor in the Amazon region and thus bring about changes in the Brazilian state's policies vis-à-vis environmental protection and the rubber tappers.

Special emphasis has been given by TAN scholars to impacts that are not easily measurable, such as the effects of transnational activism on international norms. This is certainly true in the case of scholars who adhere to a constructivist theoretical approach to the study of transnationalism. The argument is that TANs do contribute to change by shifting the norm structure of global governance. Norms, understood broadly as shared expectations held by actors about what is their appropriate behavior, are affected by the actions of TANs in at least two important ways: (1) as references by which TANs hold states and other actors accountable and (2) by the creation of new norms or the initiation of norm shifting. Either by upholding existing norms or by proposing new ones, TANs have an impact on how people in general think about the world.

Notwithstanding various successful examples of norm and policy shifting, evidence of the impacts of the campaigns launched by TANs is mixed and remains an important topic for future research. More specifically, longer term studies are needed to understand whether these changes are sustained through time.

Marisa von Bülow
Universidade de Brasília
Brasília–DF, Brazil

See also Civil Society; Social Movements

Further Readings

Christensen, I. (2008). *Transnational advocacy networks in the international system: Lessons from Ecuador*. Saarbrügen, Germany: Verlag Dr. Müller.

Keck, M., & Sikkink, K. (1998). *Activists beyond borders: Advocacy networks in international politics*. Ithaca, NY: Cornell University Press.

Khagram, S., Riker, J. V., & Sikkink, K. (Eds.). (2002). *Restructuring world politics: Transnational movements and norms*. Minneapolis: University of Minnesota Press.

Nye, J., & Keohane, R. O. (1971). Transnational relations and world politics: An introduction. *International Organizations, 25*(3), 329–349.

Rodrigues, M. G. M. (2004). *Global environmentalism and local politics: Transnational advocacy networks in Brazil, Ecuador, and India*. Albany: State University of New York Press.

Tarrow, S. (2005). *The new transnational activism*. Cambridge, UK: Cambridge University Press.

African Political Thought

The expression "African political thought" can be considered a prima facie misunderstanding in that it oversimplifies the pluralism of thoughts that has always existed in Africa. Deep divergences characterize the African ideological landscape. While Jomo Kenyatta, the Kenyan political thinker who was the first prime minister of Kenya and its president from 1964 to 1978, positioned precolonial Africa as the reference for the rebuilding of independent Africa; Léopold Sédar Senghor, president of Senegal from 1960 to 1980, idealized the encounter between Africa and Europe as the starting point of any modernity in Africa. In the same way, the call for the immediate building of a United States of Africa was contradicted by the advocacy of a step-by-step process of regional integration.

Pluralism of political thoughts is not the lone obstacle besetting the emergence of a unified school; alongside it, there is also the issue of pluralism in meaning. As the matter of fact, the notion of "African political thought" is influenced by several dialectics, including those between the African continent and the diaspora, between North Africa and sub-Saharan Africa, and between tradition and modernity. This entry first examines each of these dialectics. It then describes the substance of African political thought through two main categories: (1) "againstism" and (2) "renaissance."

Dialectics Influencing African Political Thought

The African Continent and the Black Diaspora

The first dialectic is that between the African continent and the Black diaspora. For instance, without reference to the Jamaican thinker Marcus Garvey or the American thinker William Edward Burghardt Du Bois, it is impossible to understand the genealogy of Pan-Africanism; the Pan-Africanism that constitutes the main paradigm of African political thought is initially a Diaspora idea. As strongly contesting colonization and neocolonialism, African radicalism has been elaborated mainly by the French citizen of Martinique Frantz Fanon. The relationship between these icons of African political thought and the African continent is linked to the color of their skin. African political thought is to some extent a Black political thought. It therefore appears that the idea of African political thought is linked to that of Black political thought. The debate among the diaspora in the United States, in the Americas, and in the modern world has never been considered by the continental African thinkers as foreign to their concern. Senghor and Kwame Nkrumah were always in conversation with Black American thinkers such as Du Bois and Martin Luther King Jr.

The Relationship Between Sub-Saharan Africa and North Africa

The second dialectic is that between sub-Saharan Africa and North Africa. This one concerns the racial divide between "Blacks" and "Arabs." Africa is not a continent with ethnic or racial homogeneity. The impact of this on African political thought is the cohabitation of "Pan-Africanism" and "Pan-Arabism" in the continent. Even if Pan-Africanism is unification's ideology, some of its authors restricted Pan-Africanism to "Black Africa." The Afrocentric Senegalese thinker Cheik Anta Diop established the cultural foundations of a Black African federal state. In practical terms, this idea was reformulated by President Mobutu Sese Seko of Zaire as a replacement of the Organisation of African Unity by a league of Black people and states. Seko's proposal was clearly directed against the Arabs of the continent, who were accused of being oriented more toward Middle East through the Arab League. At the same time that Nkrumah of Ghana was advocating Pan-Africanism from Cape Town to Cairo, Gamal Nasser of Egypt was defending Pan-Arabism from Rabbat to Damascus. The dialectic of the African attachment to the continent and their attraction to what is going on in Asia indicates the variable geometry of Pan-Africanism. It is worth noting that whereas Pan-Africanism and blackness are intertwined, the interaction between Pan-Arabism and Pan-Africanism is still apprehended under the prism of antagonism, instead of complementarities. The presence of Arabs in Northern Africa and the role played by Arab states such as Egypt, Algeria, and Morocco in the liberation of the continent seem to contradict the ongoing misconception of Africa as a Black continent. At this stage, South Africa's episteme of representing an African rainbow identity is a revolution. Despite the influence of African radicalism that structured the African National Congress, post-apartheid South Africa, under the leadership of Nelson Mandela, transcends racial cleavages: To be an African becomes a cosmopolitan idea that can accommodate Indians, Whites, and Blacks. It is the advent of a postracial African that was envisaged by Mandela's thought.

Tradition and Modernity

The third dialectic that operates in the realm of African political thought is that between tradition and modernity. Traditional African political thought is perceived as political thought based on what is authentic to Africa through a precolonial reference (Basil Davidson, 1998; Jomo Kenyatta, 1938). Modern African political thought is political thought generated by the colonial and postcolonial eras. From this artificial dichotomy, tradition-generated political thought is given the label of authentic African political thought. At the same time, modern political thought is stigmatized for being Westernized. This is a false opposition, because tradition and modernity are relative notions and represent intertwined realties. The search for an authentic African political thought is an impediment to the understanding of the dynamics of Africa's ideological landscape. Instead of initiating a futile debate on the criteria of authenticity, it is better to note what works as an African political thought, be it modern or traditional, Afrocentrist or influenced by "Western" ideologies.

African Political Thought as "Againstism"

Modern African political thought is to be understood through the mediation of the relation with the "intimate enemy," that is, the Western world made up of colonial powers. That is why African political thought focuses on accusation. In this regard, the radical discourse dominates the ideological landscape. The Western world is presented as the enemy on the basis of the experiences of slavery and colonialism. These two experiences are seen as part of the same trend of African humiliation in the modern world. If one agrees with Carl Schmitt that the designation of the enemy is the starting point of politics, then African thought is political and unified when African thinkers define a common enemy as in the following examples:

In the 1930s, Kenyatta accused the "Europeans" of having destroyed the human dignity of the African, denigrated his culture and religion, and stolen his land (Kenyatta, 1938). This discourse on colonization as a brutalization process is shared by the main radical thinkers.

Nkrumah's opposition to the Western world concerned its colonial and postcolonial faces. In his analysis of the European domination of Africa, Nkrumah (1961) asserts,

> The white man arrogated to himself the right to rule and to be obeyed by the non white . . . Under this cloak, the Europeans robbed the continent of vast riches and inflicted unimaginable suffering of the African people. (p. xi)

Despite the independence of African states, Nkrumah identified neocolonialism as "the last stage of imperialism." Neocolonialism is the continuity of other forms of Western domination:

> In place of colonialism . . . we have today neo-colonialism. . . . The essence of neo-colonialism is that the State which is subject to it is, in theory, independent. . . . In reality its economic system and thus its political policy is directed from outside. (Nkrumah, 1965, p. 30)

Amilcar Cabral's (1970) denunciation of colonialism is based on the fact that the Western colonial powers had taken "up arms to destroy or at least to paralyze (the) cultural life (of Africans)": Foreign domination can be maintained "only by the permanent, organized repression of the cultural life of the people concerned." Colonization and neocolonialism are explained as negations of the history and culture of African peoples.

Fanon, in his writings, formulated an absolute clash between Africa and Europe: Colonialism is presented as "fundamentally unforgivable" (Fanon, 2006, p. 118) and Europe as having proceeded with cynicism and violence against Africa.

These categories and representations with their Marxist background have exercised a real ideological hegemony in Africa and still do so today. In the intellectual domain, the conservative trend of African political thought constituted a minority. The denunciation of Europe is common to both radical and conservative African political thought. Disunity appears at the level of the relationship with the colonial power: When a conservative African political thinker such as Senghor forgives the former colonial power and regards it as a friend and partner, radical thinkers envisage struggle. The synthesis of this historic cleavage has been made in postcolonial South Africa through the "truth and reconciliation" process under which a radical liberation movement, African National Congress, accepted to cooperate with apartheid elites.

African Political Thought as a Renaissance Political Thought

African political thought is dominated by the idea of renaissance. It is a normative political thought generated by the traumatism of slavery, colonialism, and underdevelopment. These historical events constitute the implicit or explicit background of African political thought.

The renaissance from the colonial or neocolonial domination is to be obtained through struggle for national liberation. It is a cultural struggle. The liberation movement defends the culture of the people (Cabral, 1970). The struggle is also a military one implying the use of revolutionary violence (Fanon, 2002).

The renaissance following the colonial divide of Africa and its underdevelopment is to be made through the creation of United States of Africa. Despite its practical failure, this ideal still dominates African political thought (Nkrumah, 1963).

The renaissance is a terrain in which several national ideologies situated between Marxism and liberalism, such as "African socialism" or "communal liberalism" flourished.

The renaissance political thought is also a thought of liberty and equality. Under the condemnation of colonialism, apartheid, and authoritarianism, what are at stake are liberty and equality (Mandela, 1995). African political thought then appears in its universality. However, as elsewhere, African political thought is also made up of exclusion of the other. The light of renaissance is contradicted by the darkness of discourses full of hate. Ethnic exclusion and genocide in Africa have always been justified "ideologically."

Luc Sindjoun
University of Yaoundé II
Yaoundé, Cameroon

See also Colonialism

Further Readings

Cabral, A. (1970, February 20). *National liberation and culture*. Paper presented as part of the Eduardo Mondlane Lecture Series at Syracuse University, New York.
Davidson, B. (1998). *African civilization*. Trenton, NJ: Africa World Press.
Du Bois, W. E. B. (2007). *Les âmes du peuple noir* [The souls of Black folk]. Paris: La Découverte.
Fanon, F. (2002). *Les damnés de la terre* [Wretched of the earth]. Paris: La Découverte.
Fanon, F. (2006). *Pour la révolution africaine* [Toward the African revolution]. Paris: La Découverte.
Kenyatta, J. (1938). *Facing Mount Kenya*. London: Martin Secker & Warburg.
Mandela, N. (1995). *Un long chemin vers la liberté* [Long walk to freedom]. Paris: Fayard.
Nkrumah, K. (1961). *I speak of freedom: A statement of African ideology*. London: Heinemann.
Nkrumah, K. (1963). *Africa must unite*. London: Heineman.
Nkrumah, K. (1965). *Neo-colonialism, the last stage of imperialism*. London: Thomas Nelson & Sons.
Senghor, L. S. (1988). *Ce que je crois: Négritude, francité et civilisation de l'universel* [What I believe: Négritude, Frenchness, and universal civilization]. Paris: Grasset.

Agencies

In this entry, the concept of "agencies" is defined, and a short review of its history and relevance for public administration practice is provided. Finally, some key issues that have emerged from the evergrowing literature on this topic will be discussed.

Definition

Agencies are the organizational emanation of administrative decentralization in the public sector by which large monolithic core departmental bureaucracies are "unbundled" into smaller executive organizations that operate at a distance from the center of government. There are many descriptions of the administrative species "agency" in the international literature, such as "nondepartmental public bodies," "quangos," "nonmajoritarian institutions," "quasi-autonomous public bodies," and "arm's-length agencies," only to name some. For reasons of clarity, the term *agency* will be used throughout this entry. There is, however, a variation in the interpretation of the concept "agency" in different countries. Some argue that what an agency is, and what functions it does, may vary considerably across national cultures, legal systems, and political systems.

Although the definition of an agency, or the way in which "agencies" are conceptualized, may differ from country to country, there are some common features by which agencies can be characterized:

> First, in most cases, agencies are public law bodies. This means that they are subject to at least some public law procedures.
>
> Second, agencies have some capacity for autonomous decision making with respect to their internal management and/or policies.
>
> Third, agencies are structurally disaggregated from the center of the governmental apparatus (core ministry). This means that in most cases, agencies are not directly subordinate hierarchically to the oversight authorities, such as the central administration of government or the minister.
>
> Still, fourth, agencies are formally under at least some control of ministers and governmental departments, as there often is a (larger or smaller) degree of political responsibility for what the agency does.
>
> Fifth, in most cases, agencies are (at least partly) staffed with civil servants and (at least partly) financed by the state budget.

In the literature, the distinction is made between internally autonomous agencies and externally autonomous agencies. Internally autonomous agencies have a quasi-autonomous status within the central governmental department. These agencies are often labeled as "ministerial agencies," as they do not have a legal personality of their own (they are part of the legal person of the government, the ministry, or the department). This implies that the hierarchical relationship with the minister remains unchanged and that the minister remains responsible for the organization. Internally autonomous agencies do have extended managerial flexibilities, however, compared with the core governmental organizations (from which their "autonomous" status is derived). Examples of internally autonomized agencies are the Next Steps Agencies (United Kingdom [UK]), the *Agentschappen* (the Netherlands), or many federal agencies in the United States. Externally autonomous agencies are often labeled as "nonministerial agencies" because of their own legal personality and the fact that they are structurally detached from the core of the governmental apparatus. In most cases, the legal personality implies that agencies are controlled and steered by an intermediary body (like a board of governors). The existence of a board implies that the ministerial responsibility toward the agency decreases and often is only political in nature. Examples of externally autonomous agencies are the so-called nondepartmental public bodies (UK), the *zelfstandige bestuursorganen* (ZBOs; the Netherlands), or the independent administrative institutions (Japan).

Historical, Theoretical, and Empirical Context

Agencies: Not Something New at All

With the rise of new public management (NPM) as one of the dominant administrative doctrines of the past decades, many countries began reorganizing their administrative apparatus. One particular reform that could be observed in many countries was the so-called unbundling of the traditional large monolithic bureaucracies. For example, in the UK under Margaret Thatcher, the Next Steps Program dissolved the central bureaucracy in dozens of smaller executive agencies that were responsible for the implementation of public policies. This example was followed by similar initiatives in other countries such as the Netherlands or, more recently, Belgium (Flanders). However, this kind of public sector reform can be observed not only in the heartland of NPM (including many English-speaking countries) and in the broader Western world but also elsewhere. Agencification programs in developing countries (e.g., Tanzania, Jamaica), in East Asian countries (e.g., Thailand), and in Central and Eastern European countries (e.g., Latvia) have been documented extensively in the literature. Notwithstanding the fact that the 1980s and the 1990s were characterized in many countries by the establishment of many arm's-length agencies, the existence of agency-type administrative entities is not new at all. For example, for centuries the central government in Sweden has consisted of a small core bureaucracy, responsible for policy making, and many executive agencies at arm's length, responsible for implementing policies. As a second example, in the 1970s (so before NPM came into its heyday), Hong Kong's central government was unbundled into a lot of small arm's-length executive agencies. Also, the federal government of the United States has a long tradition of working with agency-like public organizations that are structurally disaggregated from the core of government.

Why Agencies?

The popularity of the agency-like administrative species in many governmental reforms is the result of the normative NPM-driven belief that this type of public organization is better suited to deliver public services in an efficient and effective way. As such, establishing agencies was considered by many governments as part of the solution in search of an answer to the crisis of the state in the 1970s and 1980s. Agencies are presumed, according to NPM advocates, to be better performers than traditional bureaucratic organizations because (a) they operate closer to the citizen/client and hence are more responsive to the client's need, (b) they are more flexible because of their high levels of managerial discretion, and (c) their (quasi-) independent status encourages them to behave more efficiently. The theoretical justification behind this—often normative—point of view can be found in rational economic theories such as property rights theory and principal–agent theory. Agencies, according to these lines of reasoning, are assumed to operate more efficiently and effectively compared with core departmental bureaucracies, because their autonomous status enables them to make their own managerial choices. The fact that agencies have the discretionary power to handle to a certain extent their own resources (personnel, finance, etc.) creates an incentive to use these resources efficiently and effectively ("let the manager manage"). On the other hand, the enlarged discretionary managerial competences on behalf of the agencies create a need to control these agencies by their principals, that is, the political and administrative oversight authorities ("make the manager manage").

In sum, according to NPM advocates, the creation of agencies is something rational, underpinned by rational-economic theories, as agencies are presumed (due to their quasi-autonomous status) to be better deliverers of public services than large, core-departmental bureaucracies. Recently, however, some scholars have argued that the agency phenomenon is more than an answer in the search for better public service delivery. These scholars argue that the agency phenomenon is a trend, a practice that is copied. There is a growing body of literature that shows that the creation of agencies, as part of public sector reform processes, is not always a rational decision. In many countries, the creation of agencies on a large scale is the result of processes of mimetic, normative, or even coercive isomorphism. In the past decades, public sector reforms that included some kind of "agencification" were considered as a best practice. Governments often tend to copy best practices from each other. Beside that many international institutions (e.g., World Bank, International Monetary Fund [IMF], and Organisation for Economic Co-operation and Development [OECD]) and multinational consulting firms have been very active in promoting NPM-like public sector reforms, including agencification programs.

Diversity

Despite the often argued fact that one is "converging" toward the agency form all over the world, which is also true to a certain extent, one should recognize that the agency form may not be that homogeneous as one might sometimes think. It can be argued that although the numerous similarities among these organizations have been created all over the world, there is also a great deal of variation, mainly due to differences in national cultures, legal systems, and political systems. The variation occurs along a number of dimensions, such as their governance structure and their relation with ministerial authorities, their degree of autonomy, and the variety of tasks they perform. In some models, agencies are organizations that are more or less directly linked to the core of government (central ministries). In English-speaking countries, such as the UK, the chief executive of the agency is responsible for an administrative oversight (a ministerial department) rather than a corporate body such as a board. This implies that the hierarchical bonds between agency and oversight are much stronger compared with the Scandinavian model, for example. In the United States, the typical agency is a public law legal person of its own, but it is still subordinate to the department within which it is located. Besides such agencies, established by a congressional act, there are also agencies that are established by and within a department. In other models, though the bonds of agencies with ministries may be weak, yet they have strong autonomy and independence. In Scandinavian countries (e.g., Sweden), agencies are responsible for implementing policies that are issued by small ministries responsible for policy formulation, although the agencies are also active in policy making themselves through their own corporate body, which has a lot of autonomy vis-à-vis the political and administrative oversight authorities. Also, in some cases, private organizations performing public tasks are considered as agencies. Governments may move some functions out of their control (e.g., through privatization), but they are still ultimately responsible for providing basic services to the public (e.g., the privatization and breakup of British Rail).

Key Issues

Roles and Functions

Agencies typically implement policies. They are, in most cases, responsible for delivering public services and public goods, such as public transport, public infrastructure, public security, and so on. This means that the role of agencies can be observed in virtually all policy fields or areas for which government is responsible. The observation that many agencies deliver public goods and services is in line with some NPM-like doctrines that advocate a so-called primacy of politics. According to such doctrines, politicians are democratically legitimized to decide on (the content and scope of) policy, while the main role of administrative actors (e.g., agencies) is to implement these policies. This strict dichotomy between policy (politicians) and operations (administrative actors) should be nuanced on at least two points. First, in reality, we can often observe in many countries that agencies have a large role to play in regulatory policy. In the policy domains of many countries, it is agencies that are the responsible actors to regulate the market in important policy fields such as energy supply, telecommunications, and postal services. These “regulatory agencies” have thus a strong policy element in their task and role (setting standards and norms, controlling market players, [helping in] defining the regulatory framework in the policy field, etc.). Second, recent research shows that agencies that are mainly involved in implementing policies (delivering public goods and services) often have a large role to play in policy making too.

Key Variables: Autonomy and Control

Agencies are generally characterized by two key variables: (1) their autonomy vis-à-vis the political and administrative oversight authorities and (2) the way in, and the extent to, which they are steered and controlled by these oversight authorities.

Autonomy

Autonomy of agencies refers to the discretion these agencies have to make decisions of their own or the extent to which the agency can decide itself about issues it finds important. In this sense, making an agency more autonomous involves the shifting of decision-making competences from external actors (e.g., the oversight authorities) to the agency itself. The level of organizational autonomy of agencies (or the extent to which the agency can make decisions of its own) is determined by the scope and the extent of these decision-making competences. Two different scopes can be discerned. First, agencies may have decision-making competences that are delegated to them from the oversight authorities concerning the choice and use of inputs or organizational resources. This is the *managerial autonomy* of agencies. An agency can have managerial autonomy with respect to several types of organizational inputs, resources such as financial resources and human resources, or other production factors such as logistics, organization, and housing and infrastructure. The second scope of organizational autonomy refers to decision-making competences of the agency in policy issues or the *policy autonomy* of the agency. The extent to which an agency enjoys policy autonomy will depend on the extent to which the agency is steered by its oversight authorities on processes, outputs, or effects. Having policy autonomy means that the agency can make decisions about the choice of policy instruments, the outputs that should be delivered by the policy, and the desirable societal effects of the policy. According to NPM-like doctrines, the ideal-typical agency has a lot of managerial autonomy, to fulfill the promise of "let the managers manage" and to ensure that the first part of the equation (more managerial competences = better organizational performance) is guaranteed. Policy autonomy, however, is low in an ideal-typical agency. As already suggested above, according to NPM, agencies are mostly expected to implement policies, not to design them (the exception may be regulatory agencies, cf. supra). In reality, however, many agencies that primarily implement policies also have a large influence in designing and evaluating policy. This may not be surprising given the fact that many agencies, as day-to-day policy-implementing actors in the policy field, are experts in their field and have a clear vision on what is happening in the policy field. A final issue concerning the autonomy of agencies is the difference between formal (legal) and real (empirical) autonomy. Formal or legal autonomy is the level of autonomy the agencies enjoy according to the laws, executive decisions, and decrees that deal with the establishment of these agencies. In these (legal) documents, it is formally described to what extent and under what circumstances agencies enjoy their own decision-making competences. In reality, however, many agencies enjoy more or less autonomy than is formally described. Research has shown that this may be due to the (political) salience of the task that the agency performs. In a politically sensitive policy field, agencies may be scrutinized more intensely by politicians compared with agencies in a politically less salient policy field.

Control

The other side of the coin is the extent to, and the way in, which agencies are controlled by their oversight authorities (which may be political—e.g., ministers—or administrative—e.g., ministries). The ways in which agencies are controlled/steered by the oversight authorities can be very diverse. First, agencies may be controlled *ex ante* ("before the facts"). *Ex ante* control is the formulation of rules and standards that give direction to the actor so that the desired output (from the viewpoint of the oversight authority) will be achieved. Oversight authorities thus control the agency beforehand. They can do so by setting general or detailed rules that constrain the autonomy of the agency in using the inputs. In this sense, *ex ante* controls take the form of authoritative mandates, rules, or regulations that specify what the agency can do. Or the oversight authority can control the agency by taking the major decisions itself or by making the decisions of the agency subject to approval by the oversight authorities (e.g., nullification rights or vetoes). Second, *ex post* control or results control is in essence checking whether the intended organizational goals have been achieved by the agency (and whether there is a need for corrective future actions). Result control involves the setting of the desired organizational results (norms), checking whether or not the organizational results meet this norm, and eventually the adaptation or continuation of the organizational practice. The intention of these "after-the-fact controls" is to motivate the agency to perform well and to make good decisions. As such, agencies are made responsible for the consequences of their decisions ("let the managers manage"). Some argue that result control can be regarded as a cycle that consists of interrelated subsystems: a planning system (setting organizational goals and norms for the agency), a monitoring system (to measure the organizational results *ex post*), and an evaluation and feedback system. The past decades, alongside the rise of the NPM, a shift toward contractualization in the controlling of agencies by their oversight authorities can be observed. In many instances, newly established agencies have performance contracts or management contracts with their oversight authorities in which the rights and duties of both parties are agreed on. Generally, the commitment of the agency is to deliver goods and services, on behalf of the oversight authorities, with a high level of efficiency, quality, and customer friendliness. Indicators are developed to be able to measure the results achieved by the agencies. The commitment of the oversight authorities generally has to do with ensuring that the agency has the necessary means (financial, personnel, infrastructural) to be able to achieve good results. Also agencies that have existed for a longer time often see their control regime shifting from *ex ante* control on inputs to *ex post* control on results and have to agree on performance or management contracts with their oversight authorities too.

Accountability and Performance

Earlier in this entry, a rather normative justification for the trend toward agencification in the public sector was discussed. NPM advocates often argue that smaller, single-task executive agencies are better vehicles to deliver effective and efficient public services and have better chances to behave innovatively and in a customer-friendly way compared with large monolithic bureaucratic organizations. The practice of performance contracts, it is often argued, also enables one to really assess whether or not agencies are performing well. However, it is not certain whether the promise of better performance has been fulfilled. Scientific evidence is at least inconclusive. One question is what "performance" is. In a public sector context, performance is a broader concept than just (economical) efficiency. No doubt that in many agencies, financial efficiency may have increased due to, among others, a better focus on the job. But other values that are of importance in a public sector context such as equal treatment and impartiality in service delivery are much harder to grasp in indicators, hence to measure for performance. The point is that increased efficiency does not always have to result automatically in increased performance on other aspects of the job that the agencies perform. Second, it is—from a methodological point of view—extremely difficult to compare performance of agencies with performance of other governmental organizations (e.g., bureaucracies). This is due to, among others, the scarcity of public functions/tasks that are in reality performed by organizations of different types in the same country (the exceptions are schools, public and private). Or the lack of null measurement of the performance of the function before it was shifted to a new organizational form (e.g., postal services before and after privatization). In sum, most commentators agree that some economical gains may be observed with agencies but that it does not necessarily imply that gains can be observed in other aspects of performance such as democracy, quality, and impartiality (which are important too, especially in the public sector).

This brings us to a final key issue to be discussed here. Performance of an organization also has to do with the accountability of that organization toward its key stakeholders. The issue of accountability is an important one because one has to be aware that in the search for the economic gains of working with agencies (efficiency; this is why they are created), there is also consideration of the possible political or societal costs because of the lack of control of these arm's-length agencies. In the case of agencies, the key stakeholders are the oversight authorities, the citizens who receive public services and goods of the agency, and ultimately, the general public or the society as a whole. Agencies are put at arm's length of the government, which makes control of the agency by government (the oversight authority) more difficult. Some authors therefore argue that the so-called vertical or upward accountability (toward the oversight authorities) may be weakened. In turn, it can be argued that agencies develop mechanisms of so-called horizontal or downward accountability, which is directed to the citizen/client of the agency (e.g., customer surveys, reporting on results, ombudspersons). Whether or not downward accountability can replace upward accountability remains a point of debate. Other observers argue that vertical accountability is large, given the fact that many agencies are subject to a multitude of control practice (nullification rights, performance contracts, audits, parliamentary hearings, etc.). In such cases, lack of accountability seems to be a lesser problem.

Bram Verschuere
University College Ghent & Ghent University
Ghent, Belgium

See also Accountability; Autonomy, Administrative; Bureaucracy; Decentralization; Delegation; New Public Management; Principal–Agent Theory

Further Readings

Bouckaert, G., & Peters, G. (2004). What is available and what is missing in the study of quangos? In

C. Pollitt & C. Talbot (Eds.), *Unbundled government: A critical analysis of the global trend to agencies, quango's and contractualisation* (pp. 3–21). London: Routledge.

Christensen, T., & Laegreid, P. (Eds.). (2006). *Autonomy and regulation: Coping with agencies in the modern state.* Cheltenham, UK: Edward Elgar.

Flinders, M., & Smith, M. (Eds.). (1999). *Quango's, accountability and reform*. Basingstoke, UK: Macmillan.

Organisation for Economic Co-operation and Development. (2002). *Distributed public governance: Agencies, authorities and other autonomous bodies*. Paris: Author.

Pollitt, C., & Talbot, C. (Eds.). (2004). *Unbundled government: A critical analysis of the global trend to agencies, quango's and contractualisation*. London: Routledge.

Pollitt, C., Talbot, C., Caulfield, J., & Smullen, A. (2004). *Agencies: How governments do things through semi-autonomous organizations.* Basingstoke, UK: Macmillan Palgrave.

Thynne, I. (2004). State organizations as agencies: An identifiable and meaningful tool of research? *Public Administration and Development, 24*(2), 1–9.

Van Thiel, S. (2000). *Quangocratization: Trends, causes and consequences*. Utrecht, Netherlands: ICS.

Verhoest, K., Peters, G., Bouckaert, G., & Verschuere, B. (2004). The study of organizational autonomy: A conceptual review. *Public Administration and Development, 24*(2), 101–118.

Agenda Setting

For any decisions to be made in the political system, an issue must be placed on the active agenda of government. In political science, agenda setting is an important topic for the analysis of collective outcomes, in particular for the evaluation of their democratic nature. Agenda setting concerns the strategic choice or selection of policy alternatives from a set of all possible outcomes on which individual, corporate, and collective actors finally decide. This entry discusses two alternative conceptions of agenda setting. One focuses more on the politics of getting an issue to the agenda. The other body of literature focuses on the impact of agendas on the outcomes of the process and on the power of agenda setters in determining outcomes. These two bodies of literature are complementary and together provide a full understanding of the role of agendas in shaping policy choices.

The Politics of Agenda Setting

In recent years, agenda-setting theory has also become a prominent approach in political communication research, which describes the transfer of important topics to the public sector for resolution, as discussed extensively by Frank Baumgartner and Bryan Jones. This portion of the literature discussing agenda setting considers the political and social mechanisms through which problems are identified and then politicized. This literature reflects the strong political role of agendas in making policy.

The agenda-setting literature also has pointed to the varieties of different agendas that exist in the public sector and the manner in which issues are moved on and off these agendas. The systemic agenda represents all those issues that are deemed appropriate for the public sector to take into account, whether or not they are being actively considered. The institutional agendas are those that are being acted on, or at least actively considered, at any point of time. Furthermore, there are important political considerations concerning which institutions constitute the foci for the political activity around any issue. Some political systems, for example, federal systems and those with active court systems, provide would-be agenda setters greater opportunities than do those with fewer points of access.

For an issue to become public policy, it must pass through several different institutions. In most democratic systems, issues must go through the legislature, and often committees within the legislature, although the source of the idea may have been the bureaucracy or the political executive, or perhaps interest groups. Then the issue, once acted on, will have to go to the bureaucracy for implementation and then perhaps to the courts for adjudication. These movements may occur as a normal part of political life but generally require some forms of political impetus for them to occur.

The agenda-setting literature has also focused on the roles of policy entrepreneurs in the process of making policy. Issues do not move on and off agendas on their own but require individual actors who foster their development and their adoption. These entrepreneurs have to be prepared when "windows of opportunity" open that permit issues that might otherwise be excluded to come to the active political agenda.

Analytic Models of Agenda Setting and Agenda Control

The second approach to agendas and agenda setting has focused on the influence that agendas and the manipulation of agendas have on the outcomes of deliberative processes. Traditionally, agendas and agenda formation have played a prominent role since the 20th century in analytical political science, which investigates the power of agenda setters in the collective decision-making process. For democratic governance, agenda-setting power can undermine the democratic process when the agenda setter has dictatorial power in manipulating the aggregation of individual preferences to collective outcomes. According to the social choice and legislative analysis literatures, a major precondition for this threat of dictatorial agenda-setting power is the instability of collective outcomes under specific democratic voting rules, which may allow an agenda setter to manipulate (or hinder from manipulating) them. Irrespective of the instability of collective outcomes, bargaining theories and logrolling approaches also emphasize the importance of agenda setting when it creates a bargaining advantage for the first mover. In other words, actors who can make the first moves in bargaining are often the ones who are able to shape the final outcomes of the process.

According to Donald Schön and Martin Rein (1994), a more optimistic interpretation comes from the deliberative literature, which stresses the framing of problems to find (stable) solutions. According to John Dryzek and Christian List (2003), deliberation has been argued to assist in aggregating individual preferences by subdividing or ordering issues in a way that induces stability of outcomes. The discussions involved in deliberative decision making enable the participants to reach solutions that are stable and reflect the underlying preferences of those participants.

According to Kenneth Joseph Arrow (1951), most theoretical work on agenda-setting power originates in social choice theory that focuses on the (in)stability of collective outcomes in democratic systems, which results from the aggregation of actors' individual preferences under specific democratic voting rules. With regard to the paradox of voting, Duncan Black (1948) and others, such as William Riker, have found that collective outcomes result from the interaction preferences of the actors involved in decision making and the structure of the agenda, that is, how policy alternatives are ordered in the collective decision-making process of democracies. This has important implications for the democratic nature of collective outcomes and the power of the agenda setter who controls the decision-making process and the voting sequence of the actors involved. Under majority rule, this power is restricted only when a median voter guarantees the stability of collective outcomes in the aggregation of actors' individual preferences.

In many decision-making situations, however, the median voter hardly exists. The existence of a median voter requires that actors with single-peaked preferences vote sincerely on an ordered single dimension. Put differently, the existence of a median voter is unlikely when preferences are not single peaked, when more than one dimension exists, or when actors vote strategically. From these restrictive conditions, under which alone the median voter theorem holds, early studies such as Black (1948) show that an agenda setter is able to achieve any outcome under majority rule. This insight on the dictatorial power of agenda setters applies more generally to all minimally democratic voting rules. Another insight from research on the (in)stability of collective outcomes is that—in policy spaces with more than one dimension—outcomes are almost always influenced by the characteristics of the decision-making processes. When instability of the collective outcome is common among the participants, democratic governance is highly threatened by dictatorial agenda setting.

This skeptical view on democratic governance has stimulated further discussion and research on the power of the agenda setters in collective decision making. More specifically, the crucial question has turned toward the extent to which an agenda setter can select and perhaps manipulate the decision-making process and, if so, what implications follow from this manipulation for collective outcomes. Some of the most influential developments in this vein of research come from studies of decision making in democratic legislatures. In these analyses, the focus is on who is able to draft and to amend legislative proposals that are finally adopted by the floor. On the most general level, this literature distinguishes between different legislative procedures by the level of institutional restrictions that define the power relationship between an agenda setter and the floor. More precisely, they find that two provisions often restrict the power of the agenda setter in democratic legislatures, the competitive right to initiate proposals and the amendment right on the floor of the legislature. In many democratic legislatures, every (individual, collective, or corporate) legislator has the right to draft a proposal, which can be adopted by the floor under either closed or open rule, the latter allowing for making amendments.

With regard to the findings in social choice theory, the power of the agenda setter is highest in legislatures without the competitive right of initiative, henceforth referred to as operating under closed rule. Under this condition, the agenda setter has complete control over drafting legislative proposals, and generally, there are limited options for amendment. The sole right of initiative is particularly decisive when the agenda setter can "keep the gate closed," that is, he or she can also ignore requests for drafting proposals or possibly for amendments. However, in some legislatures, the right of initiative is indeed in the hands of a single institutional actor to increase accountability for taking legislative action. For example, in the European Union (EU), the European Commission has the exclusive right to initiate legislation. However, scholars debate whether the Commission is capable of keeping the gates of initiative closed when the member states or the European parliaments request the European Commission to take legislative action. But regardless of the answer to this debate, the provision of an exclusive right of initiative suggests that there is a trade-off between the accountability for the decision and the level of agenda-setting power, which is sometimes minimized by distributing agenda-setting power across specialized committees. Kenneth Shepsle and Barry Weingast (1984) show that specialized committees can influence the outcome on the floor by drafting proposals. Hence, the competitive right to initiate proposals may open the gates for alternative proposals and thus reduce not only agenda-setting power but also accountability for taking action.

Another institutional provision to reduce agenda-setting power concerns the amendment right on the floor of the legislature. Under closed rules, the floor can only adopt or reject a proposal, while under open rules, it can amend the proposals. Therefore, the fundamental question is whether making amendments is subject to (additional) restrictions coming from internal legislative rules. For example, the proposals of the European Commission can formally be amended by the European Council and sometimes by the European Parliament in the following stages of the legislative process, which would suggest a lower agenda-setting power for the European Commission. However, since making amendments requires an absolute majority in the Council, and sometimes the additional bicameral consent of the European Parliament, the European Commission can exploit these higher voting hurdles in a way that makes amendments unlikely. To the extent that the European Commission is able to exploit this (bicameral) restriction, the agenda setter remains de facto powerful and can shape the final outcome.

In powerful, transformative legislatures, such as the U.S. Congress, the floor retains a substantial capacity to amend proposals coming from an agenda setter or from committees that have been preparing the legislation for consideration by the plenary body. These processes tend to minimize the powers of agenda setters and make the decision-making process less predictable than if there were greater control exercised through some agenda setter. That said, such an open-rule system may have more positive democratic effects than the more constrained agenda-setting processes.

Another feature of the amendment process is that legislators often face the problem of selecting among several amendments. As Björn Rasch (2000) notes, theoretically, a large number of voting procedures exist for selecting among several amendments, but legislatures conventionally use two procedures: the successive procedure and the elimination procedure. In the successive procedure, legislators vote successively on each amendment and decide whether or not it should be adopted. If an amendment is adopted, the process ends; if it is rejected, the legislators vote on the next amendment in the specified order. In the elimination procedure, the legislators pairwisely compare and vote on amendments. The rejected amendment is dropped, and the remaining is matched against the next amendment for a vote, and so on. Under both procedures, legislators can make amendments, but the sequence of voting (and the control thereof) can highly influence the collective outcome. As under closed rule, an agenda setter can manipulate the collective outcome by determining the voting sequence when legislators vote sincerely. In general, amendments that are introduced late on the agenda have a higher probability of adoption. However, when legislators vote strategically and anticipate the outcome, the reverse is also possible.

Although this research commonly finds that agenda setters remain powerful actors in democratic legislatures, some additional insights into the collective instability of outcomes in legislatures warrant greater empirical attention. Theoretically, the agenda setter can produce any collective outcome when legislators vote sincerely, but the scope of manipulation is often restricted when extreme outcomes are much more difficult to achieve. A major reason for this difficulty is that the voting sequence usually becomes very long and complicated when collective outcomes hardly match with the preferences of the legislative majority. In practice, as Rasch (2000) notes, such complicated sequences are hardly observed in legislatures. Another argument against the importance of the instability of collective outcomes and the resulting dictatorial power of the agenda setter criticizes the different level of behavioral assumptions that are made for the agenda setter and the legislators, that is, when the former is assumed to behave strategically while the latter is expected to vote sincerely. When the same behavioral assumption is made and legislators also behave strategically, they can anticipate an outcome from the decision-making process, which means that agenda manipulation becomes much more difficult. Under these conditions, collective outcomes are located in the uncovered set, which generally corresponds to the preferences of legislators more closely.

According to Rasch (2000), in most (democratic) legislatures the structure of the agenda and the sequence of voting are not controlled by a single actor. The responsibility for formulating the agenda is often delegated to the government of the day or to the speaker, but proposals can usually be amended by the parliament or the floor. Hence, when agenda formation is de facto made under open rules, the power of the agenda setter shrinks, and the set of feasible outcomes is restricted to the uncovered set. Given that result, most empirical research suggests that democratic institutions are designed in a way that restricts the power of the agenda setter. For example, the rules can make it possible for legislators to gather more information on their preferences in the decision-making process and thereby induce more stability by providing information. As with the deliberative processes mentioned above, the involvement of more potential decision makers can produce more stable outcomes than more hierarchically determined outcomes in legislatures.

A highly effective (but rarely observed) method of agenda setting in legislatures is to employ procedures in which the intensity of individual preferences is expressed by the number of votes assigned to an alternative (e.g., border count voting). Another way to restrict the power of the agenda setter would be the application of "backward-moving" instead of "forward-moving" agendas. In forward-moving agendas, legislators first take a vote between the status quo and a single alternative, while backward-moving agendas reverse this sequence and place the vote on the status quo on the final stage. As described by Shepsle and Weingast (1984), this procedure for establishing agendas usually limits the ability of the agenda setter to manipulate the outcome of the decision-making process vis-à-vis the status quo. Hence, empirical analyses rarely support the skeptical view on the democratic nature of legislatures when the dictatorial power of the agenda setter is only derived from an inherent instability of outcomes: Cases in which such manipulations have been observed are exceptionally rare in the real world.

Another way to study agenda setting under open rules is described by bargaining theories. In this approach, the implications of open rules often follow the idea that legislators make alternating offers to each other and decide whether they should accept a proposal for a final vote on the floor or whether they should attempt to amend it. In this bargaining process, the legislators are commonly assumed to behave strategically and to make proposals that they believe are acceptable to others involved in the process. They still would, however, only make proposals that would maximize their own benefit. A key feature of these bargaining models is that they assume that time is costly; that is, the benefits from reaching an early agreement are higher than those derived from reaching the same decision later. Hence, whenever an actor drafts a new proposal, the benefit is assumed to be lower than that derived from adopting the proposal of the earlier round, and so on. As a consequence, actors do not waste their time in infinite sequences of offers and counteroffers. This assumption is often sufficient to identify a solution from the agenda-setting process; that is, when no actor has an incentive to amend a proposal, then some generally acceptable solution has been reached. This assumption about the process implies, however, that actors can benefit from making early proposals in decision-making situations. For the same reason, this also favors patient actors, who can better afford to make counterproposals in the long run. As David Baron and John Ferejohn (1989) note, under majority rule, first movers have an additional advantage in shaping their agenda because their proposal can affect the winning coalition that results from the negotiations. Irrespective of the instability of collective outcomes, bargaining theories also support the view on the importance of agenda setting in collective decision making. Apart from institutional features, these theories demonstrate that actor-specific characteristics, such as the impatience of legislators, have an important impact on which proposals come to a vote and who shapes the final outcome in the negotiations.

These insights into agenda setting are derived from fixed policy spaces, while another line of research addresses the question on how issue linkage or package dealing can change the decision-making process and collective outcomes. According to Elmer Schattschneider (1960), from a strategic perspective, actors may seek to (de-)emphasize particular issues in order to maximize their influence on collective outcomes. With regard to agenda setting, this maximization strategy can comprise several means: First, an agenda setter may bring issues to a vote on which he or she has a winning position, creating a "bandwagon" effect. Another strategy might be to introduce new issues into a policy space that others do not sufficiently address. Hence, issues may strategically be added or removed from the agenda by the agenda setter to reshuffle majorities in collective decisions. Following this idea, issues become prominent when the agenda setter is interested in their awareness.

Conclusion

These analyses of the literature on agendas suggest that analytical political science and political communication research can come together for a future common research agenda. The communication literature emphasizes the role of entrepreneurs and their contribution in shaping agendas. The analytic perspective adds an even more strategic sense and examines how the agenda setter can manipulate the political process to produce desired policy outcomes. These two perspectives are largely complementary but too infrequently are brought together in a more comprehensive analysis of the setting and manipulation of political agendas.

Thomas König and Dirk Junge
University of Mannheim
Mannheim, Germany

See also Judicial Decision Making; Policy Process, Models of

Further Readings

Arrow, K. J. (1951). *Social choice and individual values.* New York: Wiley.

Baron, D. P., & Ferejohn, J. A. (1989). Bargaining in legislatures. *American Political Science Review, 83*(4), 1181–1206.

Black, D. (1948). *The theory of committees and elections.* Cambridge, UK: Cambridge University Press.

Crombez, C. (2003). The democratic deficit in the European Union: Much ado about nothing? *European Union Politics, 4*(1), 101–120.

Dryzek, J. S., & List, C. (2003). Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science, 33,* 1–28.

Rasch, B. E. (2000). Parliamentary floor voting procedures and agenda setting in Europe. *Legislative Studies Quarterly, 25*(1), 3–23.

Schattschneider, E. E. (1960). *The semisovereign people: A realist's view of democracy in America*. New York: Holt, Rinehart & Winston.

Schön, D. A., & Rein, M. (1994). *Frame reflection: Toward the resolution of intractable policy controversies*. New York: Basic Books.

Shepsle, K. A., & Weingast, B. R. (1984). Uncovered sets and sophisticated voting outcomes with implications for agenda institutions. *American Journal of Political Science, 28*(1), 49–74.

Tsebelis, G., & Garrett, G. (2000). Legislative politics in the European Union. *European Union Politics, 1*(1), 9–36.

Aggregate Data Analysis

Aggregate data are quantified attributes of collectivities that either relate to the body of interest as a whole (e.g., the level of democracy for countries) or have been aggregated on the basis of the properties of individual members of the collective. Aggregate data could also be negative, that is, *not individual data*, meaning that aggregate data refer to bigger entities (e.g., nations, regions, or companies) than individual data do (e.g., voters, workers, or inhabitants of some geographic area). The process of aggregation can be performed by calculating sums or various means (e.g., geometric or arithmetic mean) of the frequency distribution of individual cases. Aggregate data are predominantly secondary data; that is, researchers do not usually collect such data on their own. Another distinction refers to macro- versus microdata, with aggregate data always referring to quantities on a higher meso- or macrolevel. In this entry, types and sources of such data, major problems of reliability and comparability, and the possible applications of aggregate data are discussed.

Kinds of Aggregate Data

There are three types of aggregate data:

1. data relating to a collective as a whole (e.g., population or democracy indices, number of veto players or other characteristics of political systems);
2. aggregated individual data (e.g., the Gini Index for measuring income disparities, or the unemployment rate), which can be subdivided into two types:
   a. census data, where every single case of the whole population is measured (complete inventory count)
   b. data from samples, which should fit the context of the research question and be as representative as possible; and
3. event data, which give a frequency of events within a given period, mostly gathered from media sources (event data have become important sources for analysis, especially in international conflict research).

Basic Problems With Aggregate Data

There are some basic problems in comparative studies using aggregate data that have to be taken seriously. The first problem is *whole-nation bias*, which results from ignoring subnational variation; loss of information and reduced complexity are the consequences. Second, aggregate data are often measured in monetary units (e.g., for expenditure, import or trade volumes). For reasons of comparability between countries, the different currencies are most often translated into dollars or euros, which poses the problem of finding an adequate exchange rate for standardizing to the dollar or euro. One such standardization uses purchasing power parities, which may be misleading. Another alternative for comparing aggregated data is to standardize using a percentage of a national product measure. The choice and correct measurement of such indicators (e.g., gross domestic product [GDP] or gross national product [GNP]) are difficult and problematic; official data do not account for huge differences such as those concerning black markets or forms of subsistence income.

Intercountry comparability is also undermined by the use of different bases of calculation and a dissimilar inclusiveness of the indicators. An example makes this clear: The military budget is not identical with the budget of the Ministry of Defense. There may be a number of military expenditures (e.g., for the suppression of terrorism) that fall under the authority of other departments, such as the Ministry of the Interior. Another complication is the fact that the allocation of funds for such expenditure may vary from one country to another. The varying levels of quality in collecting data among countries can serve as a source of heteroskedasticity (i.e., a nonconstant variance where one expects a constant variance in the data). Therefore, different methods of data collection (e.g., estimation instead of more exact measurement) lead to less reliable data and, in the end, to poorer estimations in regression analysis.

Depending on which institution supplies the data (e.g., the Organisation for Economic Co-operation and Development [OECD] or the International Monetary Fund [IMF] for public finance data), huge differences for measures of the same variables are often seen. Also, revisions over time can lead to serious distortions. OECD data, especially GDP, are revised every 5 years. Apart from comparisons between countries, the comparisons between different points in time within one country can be flawed as well; for example, if statistical offices change their systems of data acquisition and calculation, poor intertemporal comparability can result. Another problem regards event data that mostly depend on media sources; in such cases, selective media coverage can lead to overrepresentation of events in countries with a highly developed media infrastructure.

A different problem in interpreting aggregate data is known as the *ecological fallacy*. It occurs when inferences drawn about individuals are based solely on data at a higher level of aggregation. For example, when analyzing referenda on European integration, it turns out that rich countries—in terms of GDP—tend to vote less in favor of Europe. By contrast, Eurobarometer statistics show that in most countries individuals with higher incomes tend to vote more pro-Europe than do those with lower income. The ecological fallacy would be to falsely conclude from the strong correlation at the highly aggregated level a correlation at the individual level.

Reasons for the Use of Aggregate Data

So why should political scientists, despite all these basic problems, still be interested in aggregate data? The answer begins on a basic level of theoretical scientific principles with a paraphrase of the well-known Kantian claim that concepts without data are empty, while percepts without concepts are blind. Data without theory also are *blind*, but theory equally needs data, because standing on its own it would be *empty*. Generally speaking, social researchers need to have access to data in order to test and develop hypotheses derived from models and theory. Both deductive and inductive approaches need data. Aggregated data should be used in particular (a) when the theory is already formulated at the macrolevel and (b) when the use of primary and individual data would be problematic in terms of availability, comparability, usability, and cost.

Time series of aggregate data are also a strong resource for detecting causal mechanisms. For example, a detailed analysis of the historical evolution of democracies, such as Samuel Huntington's "waves of democracy," would have been impossible without measurements of democracies such as those made in the Polity IV database. Another advantage of aggregated data is the possibility of making predictions on the macrolevel, for example, for economic, demographic, or political developments.

Furthermore, reasons for the use of aggregate data arise from the fact that they are especially suited for the use of statistical methods (from simple descriptive statistics to the most sophisticated methods of analysis) allowing for efficient comparisons of large spatial and social units of analysis. With the use of statistics, political scientists are also able to test reliability and to replicate other analyses quite easily.

Sources of Aggregate Data

The availability of aggregate data is important when it comes to the analysis of historical data. One such aggregate data analysis, generally considered the first, was a study of suicide rates in different countries by Émile Durkheim. Durkheim came to the conclusion that suicide is more common in predominantly Protestant countries than in Catholic ones. Max Weber was another early user of aggregated data, using information provided by the national statistical offices, which were established in the late 18th century (1796 in Sweden, 1797 in Norway, 1800 in France, 1829 in Austria). National statistical offices today are still one of the major sources for aggregate data. Other data suppliers are international organizations (e.g., UNSTAT, Eurostat, the World Bank, the IMF), which collect and prepare national data for international comparisons. One of the most comprehensive archives for international comparisons, especially for the developed countries, is hosted by the OECD. Its data archives developed rapidly (in part, due to advances in computer and Internet technology) and provide new research possibilities in terms of quantity, quality, and access to data. The steady increase in data availability gives researchers the opportunity to tackle new scientific problems and, by using data collected at different points in time, to employ longitudinal research designs. National social science data archives such as the Inter-University Consortium for Political and Social Research (ICPSR) in the United States, the UK Data Archive hosted by the University of Essex, and the "Zentralarchiv" at the University of Cologne, Germany, provide access to specific social science data, often international as well as national data. There are also special data archives serving particular purposes, such as the Luxembourg Income Studies (LIS), which provide comparable income data.

Although easy to use, these data archives bear some potential risks as well. The first risk is empiricism, or as Jan-Eric Lane (1990) put it, there is the "constant hazard that data considerations govern the conduct of comparative inquiry instead of theoretical considerations" (p. 191). Data archives are often constructed without an underlying social science theory, obstructing a sensible selection of information, which in the end could result in a boundless gathering of data that are inapplicable or unusable. This problem of too much, perhaps irrelevant, data stored in archives can be explained as most archives have some bias toward certain types of data that are readily stored; furthermore, relevant information that is more difficult to measure is thus underrepresented. Other less serious problems of data archives that can nevertheless result in a lot of time-consuming work for the social scientist are the contextual problem and the processing problem. The first concerns the fact that data in most archives have been assembled only for certain types of research problems and thus may not be suitable for other research questions—one often has to at least reassemble the data in a way appropriate for testing one's hypotheses. The latter refers to the myriad possibilities of data storage and formatting, frequently impeding the simple use of data with the statistical package of choice.

In addition to data archives, printed sources can be used to construct a data set: The classical way is to take handbooks with either thematic or regional foci. Examples of the first kind are Dieter Nohlen's compendia of elections in the world, the electoral almanac of Thomas Mackie and Richard Rose, or the *Political Data Handbook* by Jan-Erik Lane, David McKay, and Kenneth Newton. Furthermore, the revolution in information technology has cleared the way for new possibilities to gather data, so that the Internet has become the most important data source for scholars around the world. The quantity of online resources is enormous, but from an academic point of view the quality, especially in terms of the validity of the data, is mixed. Many privately run websites do not meet scientific standards.

Analyses That Can Be Done With Aggregate Data

Existing data sets can be used either for replicating former studies (for which these data sets have once been made) or for running further analyses. In a metastudy on the use of data sets in comparative politics, Andreas Schedler and Cas Mudde found that the data sets used in articles published in six of the leading English-speaking journals of political science between 1989 and 2007 show some striking characteristics:

1. Nearly half of all observed studies using data sets are single case studies and only a minority (24.5%) are large-*N* studies with more than 20 countries in the sample. Large-*N* studies, however, are becoming more and more common. In 1989, only one of the 30 studies had more than 20 cases; in 2007, these were already 41.4%.
2. Depending on data availability, most studies center on the Western world and only on a small subsample of all the thematic fields comparative politics has to offer; thus, it is a small world, the world of comparative political statistics.
3. In accordance with these findings, it is not surprising that the well-known global political data sets, such as Freedom House, Correlates of War, or Minorities at Risk, are not used as frequently as expected; rather it is the class of country-specific data sets that is most often applied.

The last important finding of their study is that only a very small minority of researchers are able to perform their analyses without modifying an existing data set. Most scientists have to modify, update, or refine existing data sets to fit their research questions. For a considerable number of studies (38.7%), comparative scholars had to completely construct new data sets on their own.

In general, each aggregate data analysis starts with a simple description of the data. This task can be performed by using either tables or graphical means such as bar charts or box plots. For some simple comparisons, this level of analysis may be sufficient, but most researchers have more specific research questions that they want to address via aggregate data. Let us assume that a political scientist is searching for the causes of government termination in parliamentary democracies; that is, the dependent variable is the time in office. The theory states that in addition to other factors, the polarization of the parliament, the economic environment, as well as the time already spent in office should play a significant role in early terminations. The first step of the analysis would be to select a sample of cases. For this research question, a most similar case's design seems appropriate, resulting in a selection of all countries having a sufficiently long record as a parliamentary democracy (say, 20 years) and having all necessary information for testing the theory available. This leaves us most likely with the OECD countries and thus quite a small number of cases (approximately $N = 20$)—that is, if we consider each country as one case. Doing so, we could compare the mean government duration of the countries with the means of the explanatory variables, for example, via bivariate cross-section regressions. The results would nevertheless not be very strong because the relationship between the independent variables and the dependent variable is likely more complex than a simple bivariate relation and the independent variables are probably intercorrelated. Therefore, it makes sense to test the theory with multivariate regression methods.

Here, a problem arises that is quite common in social science. On the one hand, there is a small N, and on the other, there are very complex models, consisting of many possible explanatory variables. This constellation generates a severe problem for statistical analyses, as the degrees of freedom are in these circumstances mostly not sufficient for a sound estimation of statistical models. There are two ways to handle this problem: (1) decrease the number of independent variables or (2) increase N. One way to minimize the explanatory variables is to construct indices. In our example, the economic environment measured by the two variables, *inflation* and *unemployment rate*, could also be covered by a combination of both, called the *misery index*. Decreasing the number of independent variables can be done only in special circumstances; in most instances, researchers will have to exclude certain explanatory variables from their models. The other possibility is to increase the number of cases. This task is generally quite difficult. A potential alternative is to not only look at the national level but also include subnational entities, when this is appropriate. In our example of government duration, this could be possible in some federal states, but the independent economic variables being defined on the national level undermine a meaningful comparison. Another possibility, which has become quite popular during the past few decades, has been the use of pooled time-series cross-section analysis. The pooling of yearly data into a matrix generates more cases (more precise observations) for the analysis. This procedure nonetheless raises new problems such as autocorrelation, which violates the assumption of random distribution of cases that is required by the standard ordinary least squares (OLS) regression. Pooled analyses, therefore, have to be used—if at all—in an extremely careful manner. When all these procedures do not result in a satisfactory cases to independent variables ratio, qualitative comparative analysis (QCA), a technique using Boolean algebra, could be applied. Fortunately, in our example, it is easier to maximize N, because every government can be regarded as a single case, and the data do not have to be aggregated further. Thus, it is possible to run multivariate regressions on a sample of all governments. However, the best methodological choice is probably not a regression but an event history model, which takes, apart from the other independent variables, elapsed time as a relevant factor into account.

Uwe Wagschal and Sebastian Jäckle
University of Heidelberg
Heidelberg, Germany

See also Qualitative Comparative Analysis; Regression; Time-Series Analysis

Further Readings

Landman, T. (2008). *Issues and methods in comparative politics: An introduction* (3rd ed.). London: Routledge.

Lane, J.-E. (1990). Data archives as an instrument for comparative research. In E. Oyen (Ed.), *Comparative methodology: Theory and practice in international social research* (pp. 187–202). London: Sage.

Lane, J.-E., McKay, D. H., & Newton, K. (1997). *Political data handbook: OECD countries.* Oxford, UK: Oxford University Press.

Mackie, T. T., & Rose, R. (1991). *The international almanac of electoral history.* London: Macmillan.

Nohlen, D. (Ed.). (2001). *Elections in Asia and the Pacific: A data handbook.* Oxford, UK: Oxford University Press.

Pennings, P., Keman, H., & Kleinnijenhuis, J. (2005). *Doing research in political science.* London: Sage.

Schedler, A., & Mudde, C. (2008, August). *The quantitative skeleton of comparative politics.* Paper presented at the 104th annual meeting of the American Political Science Association, Boston, MA.

Alienation

Political alienation was introduced in political and social thought in the 19th century. From the beginning, the concept was broadly, ambiguously, and inconsistently used by scholars, mainly because of its ideological connotations and purposes. In Hegel's and Marx's analyses of civic and capitalist society, alienation meant a state or a process of human disconnectedness. According to Claude S. Fischer, "alienation is the state in which the actor fails to perceive a positive interdependence between himself and social relationships or other objectifications" (1974, p. 18). According to Melvin Seeman, the most important elements of this conceptualization are the following:

1. Alienation refers to a kind of "discrepancy"—something that is only implicitly entailed in the definition before.
2. The individual himself or herself or his or her social environment functions as a target or a source of alienation.
3. Alienation can be regarded as a state or as a process.
4. It can be considered as a social situation/relationship, an individual mental state, a social circumstance, or a sentiment.

These aspects are discussed below.

In contemporary social science, alienation was restated as an essential part of the critique of advanced capitalism by authors like Erich Fromm, Herbert Marcuse, and Alain Touraine. Alienation, understood as an objective state or process, refers to a lack of socioeconomic resources, an underprivileged social status, an inferior position in a power relation, a lack of autonomy in labor relationships, or the degradation of objective life conditions. As a subjective feeling, it has several different meanings, such as a perceived loss of something considered valuable, a sense of inferiority, feelings of frustration, isolation, anomie, lack of control of one's own life conditions, or a perception of being manipulated by powerful others or by social structures. Drafted like this, the concept found its way in various scientific disciplines such as philosophy, psychiatry, psychology, sociology, and political science. Manifold social and psychic/attitudinal states and processes were subsumed under the heading of alienation, ranging from physical or mental illness to specific attitudes toward politics, such as political distrust or feelings of belonging to a discriminated social group (class, ethnicity, etc.). In modern political science, alienation was used more narrowly and referred to negative or critical political attitudes often leading to deviant political behavior. The topics that have been most prominent in empirical political research on alienation are analyses of various dimensions of the concept and its measurement and analyses of the impact of feelings of alienation on individuals' political behavior, particularly political apathy and participation in protest activities.

Other than the view of *alienation* as a general, diffuse disposition, the term was used as a multidimensional concept in empirical research. In this regard, Seeman was the leading scholar proposing a distinction of the subsequent six different dimensions of alienation, which were primarily understood as mental states:

1. *Powerlessness* refers to an individual's perception of low self-esteem or inferiority in political life or to a sense of low control of one's

own life conditions and the political circumstances one lives in. People feeling politically powerless think that they do not have any opportunity to make their voices heard in political life and to influence the conduct of public affairs. Some scholars distinguish between personal and political powerlessness, with the first set of attitudes describing the (lack of) ability to master one's own life and the second set referring to the perceived (lack of) capacity to exert some degree of influence on politics. The feeling of powerlessness comes quite close to the sense of political (in)efficacy or subjective (in)competence, which are much more familiar concepts in political science. It was often measured by using the ANES (American National Election Studies) standard items on political efficacy. The I-E (Internal-External Locus of Control) Scale developed by Rotter is the most common alternative measure.

2. *Meaninglessness* has to do with the individuals' sense of being able to comprehend political events, situations, and decisions and with the individuals' assessment of the impact of politics on their own lives. In short, people suffering from meaninglessness are cognitively disconnected from social and political life. This attitude resembles to a certain degree the cognitive component of powerlessness and thus is often interpreted as a facet of political (in)efficacy.

3. *Normlessness* as the third subdimension of alienation is largely synonymous to anomie and signifies the expectation that only the use of socially unapproved means will lead to the achievement of given goals. This expectation is due to the perception that social norms regulating individual and collective behavior do not perform their function any longer. Hence, political leaders do not react in a calculable and responsive way to citizens' demands. Those citizens who are well aware of this situation will rely on illegal or illegitimate forms of political behavior. In the practice of empirical research, normlessness is often equated to distrust and measured by the ANES standard items on trust in government. Alternative scales such as Dean's scale of normlessness and McClosky and Schaar's scale of anomie were also used in empirical research.

4. *Self-estrangement* seen as a failure of self-realization is the classical master theme of philosophical work on alienation and originated in the Marxian analysis of working conditions and relations. Accordingly, a lack of intrinsic fulfillment in work or an incapability of controlling one's working conditions is regarded as the core element of self-estrangement. In widening this narrow conceptualization, some scholars proposed regarding self-estrangement as an individual's engagement in activities that are not appreciated as intrinsically rewarding or are considered inauthentic. In empirical political science research, this dimension is not particularly important.

5. *Cultural estrangement* overlaps to a certain degree with normlessness and self-estrangement. It means the individual's rejection of commonly held values in society, while normlessness is seen as a rejection of more specific social norms. Again, the concept does not play a major role in empirical research, despite the important role attributed to value consensus as a source of social and political integration.

6. *Social isolation*, as the sixth subdimension of alienation, was originally seen as a lack of social relationships rather than as a mental state. In practice, however, both aspects cannot really be separated from each other. Lack or loss of social ties does not become a serious problem for individuals unless it will be perceived and assessed. Thus, social isolation signifies not only a whole set of behaviors and attitudes, a sense of being excluded or lonely, a lack of identity, a feeling of being rejected or discriminated as a member of a particular social group but also deviating behavior and lack of integration into social networks. Social isolation is used as an explanatory concept in some community studies, but it is clearly less influential in empirical research than powerlessness and normlessness.

Empirical social science research on alienation reached a point of culmination in the late 1960s and early 1970s when it was mainly embedded in two specific research contexts. The first one was the theory of mass society, which was a fashionable idea to many social scientists in that period. The second impetus came from empirical research on political protest activities and movements, which used the notions of alienation, cynicism, and subjective deprivation as the main explanatory concepts. The approach to alienation in analyses of mass society was far broader than in empirical research on political protest. As outlined by Seeman, the general research question of the former referred to the relationship of social change to political alienation, while the latter focused on various forms of behavior induced by feelings of alienation. An integration of these three sets of variables (structure, attitude, behavior) was largely missing in empirical research on alienation. An overview of the assumed links between the respective variables is outlined in Table 1.

Even if we agree to the idea of an impact of social change (particularly rapid and deep social changes) on the individual's perception of society and his or her own place in it as well as to the view of alienation as a source of deviant behavior, the specific hypotheses implied in Seeman's scheme are far from being conclusive. Political research focused on a smaller set of behaviors as well as attitudes. A first line of empirical research evolved in analyses of participation in civil rights protest movements and thus can be interpreted as an extension of classical studies of the behavioral impact of deprivation. Other research focused on the observation of an increase of political protest in affluent societies and on political protest as a means of privileged groups. The underlying assumption was that objective deprivation is not the only determinant of protest, even violent protest; negative feelings toward the political environment may also be a prominent source of protest activities. Two dimensions of alienation, powerlessness and normlessness, became particularly prominent in empirical research. According to the efficacy-distrust hypothesis, lack of powerlessness (efficacy) is a necessary condition for any type of political activity, since the powerless will generally abstain from participation irrespective of its specific form. Whether the powerful or efficacious will use protest or conventional means of exerting influence will depend on the presence or absence of normlessness (distrust). While the combination of efficacy and distrust leads to (illegal) protest, the efficacious and trusting use conventional forms of political participation such as voting, becoming active in parties, and so on. However, the empirical evidence on the efficacy-distrust hypothesis is not convincing, as is the case for the impact of alienation in general.

In assessing the state of research on political alienation, Seeman came to a rather pessimistic conclusion that still holds true today. According to him, research was often characterized by ideological convictions rather than by conceptual clarity and imbalanced by dealing mainly with powerlessness, normlessness, and social isolation and by neglecting the other components. The theoretical underpinnings of empirical research on political alienation have remained weak so far, the constructs have been used in an inconsistent manner, the measures were poor, and the empirical findings contradictory.

Oscar W. Gabriel
Universität Stuttgart
Stuttgart, Germany

See also Apathy; Efficacy, Political; Participation, Contentious; Social Movements

Table 1 Factors, Forms, and Consequences of Alienation

Contemporary Structural Trends	*Forms of Alienation*	*Behavioral Consequences*
Kinship to impersonality	Powerlessness	Political passivity
Traditional to rational forms	Meaninglessness	Wild strikes
Homogeneity to heterogeneity	Normlessness	Mass movements
Stability to mobility	Value isolation (cultural estrangement)	Ethnic prejudice
Enlargement of scale	Self-estrangement	Mental disorder
School Absenteeism	Social isolation	School absenteeism
		Low information level
		Suicide

Further Readings

Fischer, C. S. (1974). *Alienation: Trying to bridge the chasm.* New York: Wiley.

Schwartz, D. C. (2007). *Alienation and political behavior.* Chicago: Aldine.

Seeman, M. (1972). Alienation and engagement. In A. Campbell & P. E. Converse (Eds.), *The human meaning of social change* (pp. 467–527). New York: Russell Sage Foundation.

Seeman, M. (1975). Alienation studies. *Annual Review of Sociology, 1,* 91–123.

ALLIANCES

Alliances are among the most studied aspects of international cooperation and play important roles in most broad theories of international relations, particularly theories of war and peace. They are an important component of the security policies of most states in the international system today. This entry first defines and describes alliances, then discusses the motivations of states to form alliances and the effects of alliances on international outcomes.

An alliance is a formal agreement among independent states in the international system to cooperate militarily in the event of militarized conflict with outside parties. Alliances are distinguished from informal alignments by their codification in a written document; while states that share interests and tend to coordinate behavior may be termed *aligned*, only states that have made an international legal commitment to assist one another are *allied*. Alliances are distinguished from protectorates, unilateral guarantees, and colonial relationships by their joint commitments; all member states retain sovereignty and independence but commit to adjust their policies to meet allied goals. Commitments need not be symmetrical, but all alliance members must make some cooperative commitment. Finally, alliances involve promises of assistance in the event of military conflict with outside parties. Thus, they are distinguishable from other forms of military cooperation such as arms sales agreements and intelligence-sharing agreements.

Formation and Provisions of Alliances

Most alliances are formed through treaties. The specific promises made in these treaties can vary significantly. Some alliances commit the signatories to assist one another in the event a member state is attacked; these are often referred to as defense pacts. An example of a defense pact is the North Atlantic Treaty Organization (NATO). Other alliances provide for states to assist one another in accomplishing offensive goals, usually in addition to defensive promises. The Pact of Steel, signed by Germany and Italy in 1939, is an example of this sort of alliance. Many alliances, however, fall short of guaranteeing active participation in conflicts that may arise and instead commit the member states to remain neutral and provide no support for the adversary of an ally in the event the ally becomes involved in conflict or commit the members to consult in the event of threat and make every attempt to produce a coordinated response, without any specificity about what that coordinated response might be. The Molotov-Ribbentrop Pact (also known as the Nazi-Soviet Nonaggression Pact) signed in 1939 is an example of a neutrality pact, and the Treaty of Good Neighborly Friendship and Cooperation signed by Russia and China in 2001 is an example of a consultation pact (also known as an entente).

In many cases, alliance treaties also specify the conditions under which the obligations come into force. Rather than applying to any military conflict, an alliance may, for instance, apply only to conflicts with specific adversaries, in specific locations, or in conjunction with a specific dispute. The NATO treaty, for instance, specifies that it applies only to the home territories of members and does not commit members to defend their allies' colonial possessions. Defense pacts signed by France with both Poland and Czechoslovakia at Locarno, Switzerland, in 1925 commit the states to assist one another only in the event of unprovoked attack by Germany. And some treaties (for instance, the Convention of London signed by France, the United Kingdom, and Italy during the Crimean War) are limited to a specific ongoing conflict, committing the states to work together only through the conclusion of the current war.

Alliances also vary in the level of peacetime investment and coordination they require. While some alliance treaties only require action in the event conflict emerges, others provide for the development of organizations and military integration during peacetime, for one state to place troops on the territory of an ally during peacetime, and/or for shared development of defense and foreign policy. In addition, alliances vary in size and length of term. Some alliances are bilateral agreements, while others are large multilateral entities. Some alliances are formed for temporary purposes, for instance, for the duration of an ongoing war; whereas others are designed to last for long periods of time, for instance, 20 years, with provisions allowing for renewal.

Purposes of Alliances

So why do states form alliances? The most basic reason is that by joining their military forces with those of other states, they become stronger fighting forces. Given that the international system is anarchic, all bargaining among states takes place in the shadow of the use of force. States that can muster more military force can retain their security and compel concessions from others. States maintain their security by deterring threats against them, and successful deterrence requires convincing potential adversaries that pursuing gains at a state's expense through the use of force would be unsuccessful and/or prohibitively costly due to the state's ability to wage a successful war. Similarly, states can receive concessions from other states by convincing those states that resisting demands will be unsuccessful and/or prohibitively costly due to the superior military power of the challenging state or coalition. In the event that deterrence or compellence fails, and states find themselves at war, they are more likely to win the war with allied support than without.

Alternatively, states may use alliances to reduce their individual defense burdens. To the extent that economies of scale exist in the provision of defense, states might find it more efficient to combine their defense preparations with other states rather than bear the full burden of defense provisions on their own. For instance, many states found it more efficient to rely on the U.S. or the (former) USSR nuclear umbrellas during the Cold War rather than to develop their own nuclear weapons programs. Similarly, states may have different comparative advantages in defense (e.g., land power vs. sea power) and benefit from combining their efforts.

Interestingly, however, many alliances are formed between strong states and weaker states. Since weaker states usually cannot offer much increased military power to strong states (with the exception of a few weak states that occupy strategically important geographic locations), one might reasonably ask why strong states would see benefit in allying with weaker states. Two explanations (which are probably both accurate) have been proposed. One is that strong states have a self-interest in the security of some weaker states due to factors such as valuable economic relationships or rivalries with other major powers and are willing to bear the costs of defending those states. The other is that weaker states are willing to offer stronger states other advantages—for instance, foreign policy control, use of their territory and resources, or compliance on particular issues of interest to stronger states—in return for an alliance.

Implementation and Impacts of Alliances

Yet given that there is no institution to provide external enforcement of contracts in the international system, one might wonder why states believe that alliances will be reliable. How can state leaders be assured that their allies will not abandon them in the event of attack? In fact, this is a significant concern and almost assuredly prevents some potentially beneficial alliances from occurring; states should be reluctant to depend too heavily on other states to guarantee their security and when deciding to form alliances must consider the incentives that their allies will have for fulfilling their alliance commitments in the event of conflict. That being said, many allies do have incentive to work with their partners due to shared interests, and the formalization of cooperation in an alliance treaty enhances the probability of future joint action. Forming and institutionalizing an alliance are costly. Not only do states have to negotiate the agreement, they then have to implement military coordination clauses and coordinate their foreign policies to make the alliance credible. In addition, violating a previous commitment can have negative repercussions for the international reputation of a state and the domestic reputation of a leader. Thus, many scholars believe that states that invest in forming alliances have increased incentive to fulfill them in the event that alliances are invoked.

So how do alliances affect outcomes in the international system? The strongest effects of alliances are on the probability that militarized disputes occur and the probability that wars expand. Alliances affect the probability that states challenge the current status quo and make threats involving the use of military force. States with allies are less likely to find themselves challenged because potential adversaries understand that to compel compliance would involve facing a joint military effort involving the target and its allies. On the other hand, states that have allies committed to help them may be more willing to challenge the status quo and threaten the use of force in the expectation that their targets will concede their demands.

Yet, while defensive alliances may deter the initiation of disputes, when deterrence fails, states with allies may be more willing to resist a challenger's demands and take the dispute to war. If this happens, the war is unlikely to remain a bilateral affair. Alliances tend to diffuse wars beyond their initial participants and create larger, more severe conflicts. Because defensive alliances both deter disputes and make it more likely that failures of deterrence result in particularly large and severe wars, it is not easily apparent whether alliances have an overall dampening effect on international conflict.

Many scholars claim, however, that alliances have spillover effects on other kinds of international cooperation. Some argue, for instance, that allies tend to trade more with one another, that allies are more likely to settle disputes among themselves peacefully, and that institutions initially formed to support alliances become useful for a wide range of other cooperative activities as well. A large number of alliance treaties include specific provisions for nonmilitary cooperation in addition to military cooperation. Thus, alliances may have an indirect pacifying effect as well.

In conclusion, scholars today have considerable agreement on a definition for alliances, on how alliances can be distinguished from other relationships, and on the varying reasons why states form alliances. Research continues in an attempt to specify more precisely the full range of effects of alliances on international outcomes.

Brett Ashley Leeds
Rice University
Houston, Texas, United States

See also Anarchy; International Relations, Theory; Power and International Politics; Security and Defense Policy; Security Cooperation; War and Peace

Further Readings

Leeds, B. A. (2003). Do alliances deter aggression? The influence of military alliances on the initiation of militarized disputes. *American Journal of Political Science, 47,* 427–439.

Leeds, B. A., & Mattes, M. (2007). Alliance politics during the Cold War: Aberration, new world order, or continuation of history? *Conflict Management and Peace Science, 24,* 183–199.

Morrow, J. D. (2000). Alliances: Why write them down? *American Review of Political Science, 3,* 63–83.

Smith, A. (2005). Alliance formation and war. *International Studies Quarterly, 39,* 405–425.

Snyder, G. H. (1997). *Alliance politics*. Ithaca, NY: Cornell University Press.

Analysis of Variance

Analysis of variance (ANOVA) is a method for decomposing variance in a measured outcome into variance that can be explained, such as by a regression model or an experimental treatment assignment, and variance that cannot be explained, which is often attributable to random error. Using this decomposition into component sums of squares, certain test statistics can be calculated that can be used to describe the data or even justify model selection. Lab experiments have become increasingly popular in political science, and ANOVA is a useful tool for analyzing such experiments. In recent years, there have been a number of laboratory experiments on the effects of campaigning and media advertising. Nicolas Valentino, Vincent Hutchings, and Ismail White (2002), Diana Mutz and Byron Reeves (2005), and Ted Brader (2005) have all performed lab experiments that aim to determine the effect that campaigning and advertising in media have on voters views and decisions. All three experiments employ an ANOVA to control for observable characteristics, interactions between treatment regimes, and significance of the relative effectiveness of treatments. This entry discusses ANOVA and its applications in greater detail.

In the familiar regression context, the "sum of squares" (*SS*) can be decomposed as follows. Assuming that Y_i is individual *i*'s outcome, $\bar{Y}$ is the mean of the outcomes, $\hat{Y}_i$ is individual *i*'s fitted value based on the ordinary least squares (OLS) estimates, and e_i is the resulting residual,

$$\sum_{i=1}^{N} (Y_i - \bar{Y})^2 = \sum_{i=1}^{N} (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^{N} (e_i^2),$$

where

$$SS_{\text{total}} = \sum_{i=1}^{N} (Y_i - \bar{Y})^2$$

is the total sum of squares,

$$SS_{\text{regression}} = \sum_{i=1}^{N} (\hat{Y}_i - \bar{Y})^2$$

refers to the variance explained by the regression, and

$$SS_{\text{error}} = \sum_{i=1}^{N} (e_i^2)$$

is the variance due to the error term, also known as the unexplained variance. Commonly, we would write this decomposition as

$$SS_{\text{total}} = SS_{\text{regression}} + SS_{\text{error}}.$$

The equations above show how the total variance in the observations can be decomposed into variance that can be explained by the regression equation and variance that can be attributed to the random error term in the regression model.

ANOVA is not restricted to use with regression models. The concept of decomposing variance can be applied to other models of data, such as an experimental model. The following is the decomposition of a one-way layout experimental design in which an experimenter randomly assigns observations to one of *I* treatment assignments. Each treatment assignment has *J* observations assigned to it. In the case of a randomized controlled trial with only one treatment regime and *N* subjects randomly assigned to treatment with half a probability, this would mean that $I = 2$, one treated group and one control group, where each group has size *J*. In this framework, the variance decomposition would be as follows:

$$\sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - \bar{Y})^2 - \sum_{j=1}^{J} \sum_{i=1}^{I} (\bar{Y}_{ij} - \bar{Y})^2 + \sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - \bar{Y}_i)^2,$$

where

$$\bar{Y}_i = \frac{1}{J} \sum_{j=1}^{J} Y_{ij}$$

is defined as the average response under the *I*th treatment and

$$\bar{Y} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}$$

is defined as the overall average of all observations, regardless of treatment assignment. Commonly, this sum of squares expression is written as

$$SS_{\text{total}} = SS_{\text{between}} + SS_{\text{within}},$$

where SS_{between} refers to the part of the variance that can be attributed to the different treatment assignments and SS_{within} refers to the variance that can be described by the random error within a treatment assignment. From this, we can see that SS_{between} and $SS_{\text{regression}}$, from the regression framework, both refer to the explained variance. SS_{within} and SS_{error} both refer to the unexplained variance.

Typically an ANOVA table is constructed to describe the variance decomposition. In the regression context, where *p* is defined as the number of independent regressors and *n* is the number of observations, the ANOVA table typically looks like this:

Source	*df*	*SS*	*MS*	*F*
Regression	P	$SS_{\text{regression}}$	$SS_{\text{regression}}/p$	$\dfrac{SS_{\text{regression}}/p}{SS_{\text{total}}/n-1}$
Error	$n-p-1$	SS_{error}	$SS_{\text{error}}/n-p-1$	
Total	$n-1$	SS_{total}		

This table gives us an idea as to how to break down our analysis. The column *df* refers to the degrees of freedom. In the regression framework, the degrees of freedom for the regression is the number of parameters in the regression equation. The degrees of freedom for SS_{error} is $n - p - 1$. The column *MS* refers to the mean squared error, which is defined as *SS/df* for each row in the table. If we were in the experimental one-way layout design, then the first row would refer to the between variance, and the number of treatments less one, that is, $(I - 1)$ for I treatments. The degrees of freedom for error is defined as the number of treatments times the number of trials in each treatment less one, or $I(J - 1)$ for J trials in each treatment.

The above table also contains a column called F, which refers to the F-test. The F-test is a way of using ANOVA to determine if all the regressors in a regression equation are jointly zero. In a one-way experimental analysis, the F-test determines if the means of the treatment groups are significantly different. If we have more than one treatment and one control group, then the F-test is applied to see if any treatments are significantly different from zero. The null for an F-test is that all coefficients in our model are jointly, not statistically, distinguishable from zero. The F-test is defined as follows:

$$F = \frac{SS_{\text{regression}}/p}{SS_{\text{error}}/n-p-1}.$$

Using the ANOVA table, we can also determine the R^2 value of our treatment or model. We define R^2 as follows:

$$R^2 = \frac{SS_{\text{regression}}}{SS_{\text{total}}} - 1 - \frac{SS_{\text{error}}}{SS_{\text{total}}}.$$

The R^2 refers to how much of the variance is explained by the model. A high R^2 value means that much of the variation is explained by the model, implying that the model fits the data well and that little of the variance is explained by the random error term. In the experimental framework, a high R^2 value means that much of the variation is explained by the treatment assignment, and little of the variance is due to random error within those treatment assignments.

Model Selection and Analysis of Variance

ANOVA is generally used with linear regression to assess model selection. When selecting the best model, we seek to strike a balance between goodness of fit and parsimony. If two models fit the data equally well, the model selected should include only those explanatory variables that explain a significant degree of the variance in the response variable. The question is how to distinguish between important and trivial variables in a way that is systematic. ANOVA is one method for identifying the parameters of interest. It is important to stress that ANOVA makes all the assumptions made by normal linear regression. Furthermore, in general applications of ANOVA, all the explanatory variables must be mutually orthogonal, although in some limited cases this orthogonality is not necessary to make a reasonable justification for model choice. To determine which covariates are important for the regression model, ANOVA can be run multiple times in succession to determine if adding an additional covariate contributes any more to the explained variance.

Consider two normal linear models:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \varepsilon, \quad (1)$$

$$y = \alpha + \beta_1 x_1 + \varepsilon, \quad (2)$$

where x_1 and x_2 are explanatory variables, α is the intercept, and β_1 and β_2 are the parameters of interest. The second is obviously a simpler version of the first. We can think of Model 2 as the version of Model 1, in which β_2 is restricted to zero. For this reason, we often refer to Model 1 as the unrestricted model and Model 2 as the restricted model. The question is which model is better. If the restricted model fits the data equally well, then adding complexity does not improve the accuracy of the estimation significantly and the simpler (restricted) model is preferable.

In model selection, ANOVA analyzes the degree to which residual variance changes with the addition of explanatory variables to the basic model. Note that the vector of residuals for the restricted model (where $\beta_2 = 0$) can be broken into two components:

$$\begin{gathered} y - \hat{y}_1 = \{y - \hat{y}_2\} + \{\hat{y}_2 - \hat{y}_1\} \\ \hat{y}_1 = \hat{\alpha} + \hat{\beta}_1 x_1 \qquad \hat{y}_1 = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 \end{gathered}.$$

Thus, the vector of residuals for the restricted model consists of a vector of the residuals for the unrestricted model plus the residual difference between the two models. By construction of OLS, the vectors $(y - \hat{y}_2)$ and $(\hat{y}_2 - \hat{y}_1)$ are orthogonal, and Pythagoras's theorem implies that the sum of squares for the restricted model is just the sum of squares for the unrestricted model plus the difference in the sum of squares for the two models, or equivalently,

$$SS_{\beta_1} = SS_{\beta_1,\beta_2} + \{SS_{\beta_1} - SS_{\beta_1,\beta_2}\}.$$

While adding complexity reduces the amount of unexplained variance in the residuals, it also reduces the degrees of freedom. This trade-off motivates the principal of parsimony in model selection. Under the assumptions of OLS, SS_{β_1} and SS_{β_1,β_2} are mutually independent and have a chi-square distribution. The F-test is therefore the appropriate test to determine whether the degree to which inclusion of each additional explanatory variable in the model improves the precision of estimation. In this case, the F-test would look as follows:

$$F = \frac{\{SS_{\beta_1} - SS_{\beta_1,\beta_2}\}/p - q}{SS_{\beta_1,\beta_2}/(p - q - 1)},$$

where p and q represent the number of parameters in the unrestricted and restricted model, respectively (excluding the intercept). Under the null hypothesis, the unrestricted model does not provide a significantly better fit than the restricted model; reject the null hypothesis if the F calculated from the data is greater than the critical value of the F distribution with $(p - q, n - p)$ degrees of freedom. The models, their sum of squares, mean square, and F-test can be displayed in an ANOVA table:

Fitted Model	*df*	ΔSS	*MS*	*F*
$\hat{y}_1 = \hat{\alpha}$	1	$SS_{\text{error}} - SS_{\hat{\alpha}}$	$\frac{SS_{\text{error}} - SS_{\hat{\alpha}}}{n - (n-1)}$	$\frac{SS_{\text{error}} - SS_{\hat{\alpha}}}{SS_{\hat{\alpha}}/(n-1)}$
$\hat{y}_1 - \hat{\alpha} + \hat{\beta}_1 x_1$	2	$SS_{\hat{\alpha}} - SS_{\hat{\beta}_1}$	$\frac{SS_{\hat{\alpha}} - SS_{\hat{\beta}_1}}{n - 1 - (n-2)}$	$\frac{SS_{\hat{\alpha}} - SS_{\hat{\beta}_1}}{SS_{\hat{\beta}_1}/(n-2)}$
$\hat{y}_1 = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$	3	$SS_{\hat{\beta}_1} - SS_{\hat{\beta}_1,\hat{\beta}_2}$	$\frac{SS_{\hat{\beta}_1} - SS_{\hat{\beta}_1,\hat{\beta}_2}}{n - 2 - (n-3)}$	$\frac{SS_{\hat{\beta}_1} - SS_{\hat{\beta}_1,\hat{\beta}_2}}{SS_{\hat{\beta}_1,\hat{\beta}_2}/(n-3)}$
Total	4	SS_{error}	$SS_{\text{error}}/(n-3)$	

Often the results of an ANOVA table are used to justify the inclusion of each individual variable in the model. As the model expands from p to $p + 1$ explanatory variables, the F-test evaluates the hypothesis that the parameters $\beta_{p+1} = 0$, given the assumptions of the model are satisfied. If the explanatory variables that constitute the design matrix are all mutually orthogonal and we have the correct model, then the ANOVA results can be used to determine whether the inclusion of x_{p+1} significantly increases the fitness of the model. Without orthogonality, however, we do not know if the order in which the variables are added matters. As successive variables are added from the model, only the variance of the part of the variable that is orthogonal to the previously included variables in the model is removed from the variance of the error.

This entry has discussed the definition of ANOVA and how it is often applied to regression and experimental data. ANOVA is a decomposition of variance into component parts. There is variance that is attributable to a model, such as a regression model or an experimental treatment, and variance that is attributable to random error. ANOVA is often used to construct an ANOVA table, which succinctly presents the variance decomposition. This method can also be used to justify regression model selection where the goal is to find parsimony between fit and degrees of freedom. ANOVA can be used to determine how much extra variance a marginal explanatory variable explains while also weighing the loss of a degree of freedom. It is important to note, however, that the order in which variables are added to a model is important in these tests unless the variables are orthogonal to one another. The decomposition of variance using the analysis of variance is a powerful tool for describing data and the fit of a model.

Erin Hartman and Adrienne Hosek
University of California, Berkeley
Berkeley, California, United States

See also Quantitative Methods, Basic Assumptions; Regression

Further Readings

Brader, T. (2005). Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions. *American Journal of Political Science, 49*(2), 388–405.

Davison, A. C. (2003). *Statistical models*. New York: Cambridge University Press.

Hill, R. C., Judge, G. G., & Griffiths, W. E. (2001). *Undergraduate econometrics*. New York: Wiley.

Mutz, D., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, *99*(1), 1–15.

Rice, J. A. (1995). *Mathematical statistics and data analysis*. Belmont, CA: Duxbury Press.

Valentino, N. A., Hutchings, V. L., & White, I. K. (2002). Cues that matter: How political ads prime racial attitudes during campaigns. *American Political Science Review, 96*(1), 75–90.

Analytic Narratives: Applications

Often derided by colleagues for "possessing an *N* of 1," those who focus on case studies—Germany between the wars, France during the revolution, or Kasumpa village (see below)—nonetheless aspire to be social scientists. While their commitment to fieldwork, archival research, and qualitative methods runs deep, so too does their commitment to science. Analytic narratives offer a means for reconciling this apparent contradiction. Avner Greif, Margaret Levi, Jean-Lauran Rosenthal, Barry Weingast, and Robert Bates formulated the approach while focusing on historical cases. This entry illustrates the approach by reviewing its use in a study of Kasumpa village in the Luapula Valley of Zambia.

In this essay, "fieldwork" is defined as immersion in the lives of people who remain resident in their own social setting, while observing and recording their behavior and discussing with them their actions and the values and beliefs that shape it. This also includes, of course, knowledge of the local language. "Science" refers to the attempt to derive valid explanations, which implies two things: An explanation is valid if it follows logically from its premises, and it is valid insofar as it withstands efforts to refute it through the systematic collection and analysis of data. The best test of logical validity is formalization; that of empirical validity is the use of rigorous methods.

Taken together, these clarifications highlight a key feature of the agenda that underlies this method: the treatment of qualitative research, formal theory, and empirical methods as complementary rather than rival approaches to social research.

Illustration

Rural Responses to Industrialization: A Study of Village Zambia by Bates reports on a study of Kasumpa village, a settlement in the Luapula Valley that borders Zambia and Congo and forms a portion of the hinterland of the great mining centers of those two countries. The author of this study notes how, in the initial stages of his fieldwork, he sought to discern the values held by the residents of the village, their perception of the alternatives available to them, and the beliefs that shaped their behavior. The residents of Kasumpa village, he learned, quite rightly viewed themselves as poor and the mining towns as prosperous. Economic life revolved around the export of produce and labor to the mining towns; political life centered on efforts to persuade the government to generate money and jobs by investing in projects in the region. The residents sought to use the markets for produce and labor to gain access to the wealth of the mining towns, and they used political protest to levy a portion of it for themselves.

Based on these insights, the author then turned to theory. In expositing this work, this essay focuses on his treatment of migration, a subject long central to research in the region. The dominant theory of the time, developed by John Harris and Michael Todaro, viewed migration as a decision made by individuals as an effort to maximize their expected incomes. Several implications flowed from that formulation, and the author randomly selected 100 or so out of around 1,000 villagers to test them. If the theory were correct, more men than women would migrate to seek work in town, because there were more jobs for men in the mines, and the probability of securing a job was therefore greater for males than for females. The likelihood of departure from the village should be higher for working-age men: They were more likely to secure jobs than youths were and could amortize the costs of job searches better than old men could. In addition, according to this theory, when people left, they should journey to towns where others had settled before them, thus lowering the costs of job searches and increasing the return from migrating. The demographic structure of the village should therefore be distorted, with a gap where working-age men could normally be expected and a disproportionate number of females and old people of both sexes. An implication for politics also followed: While the migrants might choose to "exit" as a way of escaping poverty, those who remained should "give voice," pressuring the government to render fishing and maize growing more profitable. Phrased another way, while migrants might employ the market for labor as an alternative to political action, those who remained in the village should treat the market for commodities as a complement to it.

The data lend support to each implication, with young men flocking to the nearest mines, women and older men tending to remain behind, and fishermen and farmers dominating the local party branch and employing their positions to lobby the government for loans to farmers and investments in the local fisheries. However, while collecting the data, the researcher gained a deeper understanding of the migration decision—one that suggested that, in fact, the expected income model was wrong. Migration, it became clear, involved not only those who chose to migrate but also those they left behind and, in particular, their elder kin. Nor was migration the outcome of a choice made at one point of time; rather, it was a portion of an intertemporal sequence of moves in which the older generation paid for the costs of the movement of younger, working-age males from the village to town. Not only did they pay for transport and advance funds for sustenance during the initial period of job search, but they also paid for the costs of education so that the youths might better compete for jobs in town. It was made clear that they did this in the expectation of later receiving remittances from the young, thus assuring them of financial support in old age. In effect, rather than acting as individual agents, the residents of Kasumpa village acted as members of families and organized the relationship between generations so as to invest in the formation of human capital. By responding thus to the opportunities offered by the market for labor, they sought to extract income from the towns.

Stepping back for a moment, note the sequence and the interplay between different research methods. First, there was immersion and the gathering of qualitative information, then the formulation of theory. Next, the theory was tested, using survey methods and engaging further in "soaking and poking"—conversations about the relationships between parents and children, discussions about family life in the United States and Luapula Valley, and gossip about why X could live in comfort ("great kids") whereas Y lived in poverty ("the way he neglected his children, no wonder"). Then came the realization that the theory was wrong. The resultant crisis then restarted the cycle but not at a random starting point. Rather, because the failure took place within the context of a theory, it highlighted key features that had been missed the first time around.

It was clear that the decision to migrate did not result from a comparison between rural and urban wage rates, made by individuals who sought to maximize expected incomes, as the Harris-Todaro model might suggest. Rather, it resulted from the strategies used by members of families who sought to invest by forming human capital. Indeed, data from the surveys confirmed that the flow of finances between generations yielded a reasonable rate of return, for plausible rates of discount. But, the researcher realized, if this revised vision of the behavior of migrants was correct, then there was a *lot* he did not understand: According to the current theory, what was being observed, in fact, should not exist. Less dramatically, it suggested that the arrangement was highly vulnerable and that its persistence suggested that additional factors, not yet described, must be at play.

The theoretical difficulty came from two features of the behavior of the families. The first was that it involved an exchange; the second was that the exchange took place over time. The elders expended resources on the young at one point in time in the expectation that the young would later repay. But the preferences of the young should vary over time: They should be initially willing to pay but reluctant to do so when the bill came due later. Given this variation in preferences, the elders, anticipating the later actions of the youths, should not be willing to invest in the first place. The (implicit) contract was therefore vulnerable to the problem of "time consistency." And yet the elders were observed investing in junior kin.

Highlighting the anomalous nature of the finding was that local conditions reinforced the strength of the temptations facing the young. There was a long lag between the time when the elders bestowed resources on them and when they could be expected to remit money from town; the distance between the village and the town was great as well, making it easy for the young to dissimulate, claiming that the loss of a job, illness, or unforeseen expenses in town made it impossible to fulfill their obligations to village kin. The elders were of course aware of the impact of time and distance on the ability of the young to elude their obligations. And yet the elders continued to use their families in order to invest. The author of the study was thus driven by the failure of his theory to engage once again in the lives of the people of Kasumpa and to seek out what was being missed—accounting for their conduct, which otherwise appeared so rational.

Once again, theory had failed. However, once again, the failure proved generative, for it forced a search for features of village life that had been overlooked thus far. One feature was that the young hoped to return to the villages and depended on their families when seeking to do so. To retire in the village, a worker would need land on which to build a house and plant a garden. To secure that land, he would need to be *kumwesu*—that is, "one of us." To establish his bona fides, he would have to rely on his family in the village to testify to his roots there; to his participation in its affairs, as by returning for weddings and funerals; and to his continued support of the welfare of its citizens, as by remitting funds from town. Thus, the ability of the young to fulfill their plans rested on the willingness of their family to confirm that they had fulfilled their obligations to their elder kin.

Theory had therefore failed, or at least been exposed as inconsistent with the known data, and its failure had motivated a renewed pursuit of qualitative data—one focused on a search for features of reality that would fill gaps in the argument. Phrased differently, in the interplay between theory and qualitative research that characterizes the generation of analytic narratives, formal theory—which is often viewed as a deductive method—is instead used inductively.

Thus far, more emphasis has been given to fieldwork and theory than to the use of empirical methods. Recall, however, their use when testing the implications derived from the comparative statics of the expected income theory of migration. As illustrated in *Rural Responses to Industrialization*, they play an additional role as well: They are used to rule out alternative explanations.

Returning to the contributions of the elders to the young, someone might offer an alternative explanation: The elders could be viewed as altruistic rather than self-interested. Their payments to the young could be treated as gifts made out of affection. Given that both accounts are consistent with what has been observed, researchers, when marshalling empirical methods to choose between the two accounts, need to bring them to bear on data drawn from outside of the original sample. To gather such data, the author shifted his attention to the plateaus that lay above the valley. Because of the tsetse flies, the villagers in Luapula could not raise livestock; because the plateaus were free of tsetse flies, those who lived there could. There, the elders invested in cattle. Because cattle reproduce, herds grow over time, and as the urban centers grow, so does the demand for meat. All else being equal, cattle herds therefore increase in value, enabling their owners to recruit clients, bargain for brides, and bid for favors. If the elders were motivated by self-interest, then, rather than investing in their children, they would invest in their cattle; if they were motivated by affection, they would invest in junior kin. The evidence showed clearly that they invested in their herds. They stinted on the payment of school fees, so the level of education remained low. The rate of migration of cattle-keeping peoples remained low as well. Gathering empirical data from out of the original sample thus lend credence to the researcher's interpretation of the behavior of the villagers. Had affection dominated self-interest, then the flow of resources between generations on the plateau would have resembled that in the valley. Clearly, it did not. The analysis of theoretically relevant data, collected outside of the original sample, thus dispensed with an alternative interpretation, leaving the original intact.

Conclusion

The end product of this research was a "village study." But it was also a study that was the product of repeated efforts at falsification. When ideas were rendered coherent, then data were gathered to test them; when reflection exposed weaknesses in interpretation, then additional field research filled in the gap. The resultant study was not "large *N*." Nor was it axiomatic and deductive. But, as it was shaped by theory and statistical methods, the study "scaled up." It confirmed that villagers in the middle of Africa were rational and self-interested in their pursuit of wealth and their use of power, a controversial claim at the time and one still contested by cultural anthropologists. As in the treatment of the behavior of the villagers on the plateau, it supported those who contest the premise of cultural conservatism among cattle-keeping peoples. It helped redefine the understanding of how migration takes place and how rural people react to urban opportunities. And, along with the contributions of others, it helped alter the manner in which political scientists approach the study of agrarian politics.

There are some in political science who champion the "deep reading" of texts and "immersion" in rich, qualitative data. Others label themselves as formal theorists and generate rigorous proofs. And still others consider themselves to be methodologists and specialize on the development and use of systematic empirical methods. Those who develop analytic narratives view these approaches as complementary. The first yields *Verstehen* or comprehension; the second generates conviction, demonstrating that the explanation is logically consistent; and the third determines whether it is compelling—that is, it can be demonstrated convincingly to others. By combining these approaches, those who seek to produce analytic narratives "do science" even while focusing on unique events or particular settings.

Robert H. Bates
Harvard University
Cambridge, Massachusetts, United States

See also Analytic Narratives: The Method; Case Studies; Migration; Mixed Methods; Participant Observation

Further Readings

Bates, R. H. (1976). *Rural responses to industrialization: A study of Village Zambia*. New Haven, CT: Yale University Press.

Bates, R. H., Greif, A., Levi, M., Rosenthal, J.-L., & Weingast, B. (1998). *Analytic narratives*. Princeton, NJ: Princeton University Press.

Epstein, A. L. (1958). *Politics in an urban African community*. Manchester, UK: Manchester University Press.

Harris, J., & Todaro, M. (1970). Migration, unemployment, and development: A two sector analysis. *American Economic Review, 60*(1), 126–142.

Herskovitz, M. J. (1926). *The cattle complex in East Africa*. Washington, DC: American Anthropological Association.
Mitchell, J. C. (1956). *The Kalela dance* (Rhodes Livingstone Papers No. 27). Manchester, UK: Manchester University Press.
Popkin, S. L. (1979). *The rational peasant*. Berkeley: University of California Press.
Scott, J. C. (1976). *The moral economy of the peasant*. New Haven, CT: Yale University Press.
Stark, O., & Bloom, D. (1985). The new economics of labor migration. *American Economic Review, 75*(2), 173–178.

Analytic Narratives: The Method

An analytic narrative is a methodological proposal to bridge rational choice modeling with more traditional narrative explanations of phenomena in the social sciences. Typically, the methodology combines models from game theory with historical narrative. The analytic narrative project, however, does not represent a methodological breakthrough in itself. Rather, the project makes explicit the methodology that numerous scholars have adopted when combining historical and comparative research with rational choice models. Influenced by the work of Douglass C. North, the analytic narrative project started off with a commitment to understand institutional formation and change. In their book *Analytic Narratives* (1998), Robert Bates, Avner Greif, Margaret Levi, Jean-Lauran Rosenthal, and Barry Weingast offer the first systematic outline of the key elements of the methodology, and they do so by exploring institutional change in a wide range of places and times. The authors' intent, as well as that of the majority of analytic narrativists, is to investigate enduring questions of political economy, such as political order, political and economic governance, and intra- or interstate relations.

Moreover, a distinguishing feature of the analytic narrative project is that it underscores the importance of the institutional context within which historical events occur. Understanding the institutional context helps account for how and why certain events may happen. The importance afforded to the historical and institutional detail is in fact what sets analytic narrativists apart from most rational choice scholars. Both deduce their hypotheses from the assumptions of rational choice and the logic of game theory. Rational choice scholars, however, start off from a general model and then test their hypothesis with appropriate data, while analytic narrativists formulate and refine the model itself in interplay with the context-specific institutional elements of the historical narrative. A major objective of analytic narratives, shared certainly with social science research, is to successfully identify the causal factors that explain a particular historical phenomenon—that is, to determine how some combination of causes brought it about. The notion of "mechanism" is commonly used. This entry discusses the conditions under which analytic narratives are most successful in providing explanations for historical phenomena. It first examines the distinctive roles of formal theory and context-specific evidence in developing a conjecture and ascribing causality. Second, it explores the behavioral assumptions underlying strategic game theory and their implications for analytic narratives.

Theoretical Model Versus Narrative

The construction of an analytic narrative proceeds, roughly, as follows. First, the scholar acquires in-depth knowledge about the historical phenomenon of interest; that is, a detailed account of the context and the historical process, based on studying the past through primary sources or reading the already existing historical accounts. A detailed account is essential to isolate the relevant strategic elements in the interaction: the key actors, their goals, and the rules that structure their behavior. These elements can then be formalized in a model. The formulation of the model—generally a game theory model—specifies the choices, constraints, and trade-offs the actors face in the phenomenon in question. The model is supplemented with a narrative that provides a rich explanation of the meaning actors attach to their actions, circumstances, and surroundings: the significance of the local culture. The outcomes predicted by the theoretical analysis are then confronted with the narrative; the narrative serves to assess the predictions and arbitrate among possible explanations in instances of observational equivalence. Further refinement of the model, and collection of more historical detail, can result from additional iterations between the analytics and the history. Analytic narratives stand on a careful balance between context-specific detail and rigorous analytic techniques. An analytic narrative, however, is problem driven, not theory driven. Thus, the explanation of a particular event is what motivates each study. Nonetheless, even if driven by a particular case, analytic narratives are informed by theoretical modeling. The exercise of formalization, by isolating the relevant strategic elements in an interaction, helps identify the key actors and the combination of causes that can explain the strategic situation in question.

A major reason for applying models to narratives is to attain logically consistent explanations; thus, the predictions of the model must follow deductively. However, it must be pointed out that the model does not need to do the bulk of the explanation. The theory highlights the issues to be explored and the general considerations and evidence that need to be examined, while the knowledge of the historical context is used to develop a conjecture regarding the relevant institution. The theory provides categories and a framework in terms of which the conjecture about the causal mechanism is formulated. The model, thus, *constrains* but does not determine the conjecture.

The conjecture is then evaluated, refined, or even overhauled through the interactive use of a context-specific model and the historical narrative. This empirical method thereby recognizes and takes advantage of the context specificity and historical contingency of institutional analysis. A successful explanation requires a well-confirmed causal claim about why and how a certain outcome obtained—this can be done even if it is primarily the narrative rather than the model that accounts for the explanation. The explanation that the model points to should also survive competition with other explanations. These can be existing explanations that the author may be contesting or other potential explanations that could account for the historical facts. The proposed explanation must be confronted with these other explanations and shown to be superior.

Analytic narratives, by providing a detailed account of the context and the historical process, acknowledge the uniqueness of situations that take place in particular moments and at specific places. At the same time, analytic narratives seek to identify causal mechanisms that are generalizable to other situations, inasmuch as they explain a social phenomenon by identifying how some combination of causes brought it about. Analytic narratives seek to capture the uniqueness of the situation under investigation by means of the narrative while using the model to capture the general features of the type of phenomena under which this situation falls. The use of rational choice in analytic narratives, by carefully identifying the causal mechanisms, makes possible the application of the logic of one setting to another.

Explanations that can account for different phenomena using few arguments—that is, generalizable explanations—are of course desirable. Generalizability, however, is not a requirement of explanation, since a causal mechanism may be unique to a phenomenon. Thus, for an analytic narrative to provide a successful explanation, it is not necessary that the causal mechanism be generalizable. It suffices that the events the model identifies as causes and effects actually take place. If a generalizable statement of the mechanism identified by an analytic narrative is available, this is of course a plus. In short, for analytic narratives to offer successful explanations, the formalization at the heart of the narrative must correctly identify some of the causes, effects, and enabling conditions in question. This does not require that the causal mechanism be generalizable. What is required is that the events the model identifies as causes and effects actually take place and that the statement of causal mechanism survives competition with other explanations.

Behavioral Assumptions

The reliance of analytic narratives on explicit formal theorizing compels scholars to take a systematic approach and to reveal the reasoning and assumptions behind their explanations. However, by emphasizing certain aspects of reality, this reliance on models, and specifically game theory models, highlights certain scope conditions for which the analytic narrative project is better suited. The scope conditions refer to the situations in which game theory is the best tool to explain human behavior. In terms of the historical phenomena of interest and the types of questions addressed, analytic narratives are a better tool when causal explanation is found in strategic interactions. They theorize strategic interaction and not the structural conditions under which the interaction takes place. Macrolevel structural factors are taken as exogenous during modeling interactions, which implies that changes in such factors need to be incorporated as moves by "nature" and are not treated in an analytic fashion. This emphasis on strategic interaction, however, is not a weakness of analytic narratives as much as a necessity given their focus on microlevel historical data.

The analytical framework of classical game theory rests on seemingly unrealistic assumptions about the cognition, information, and rationality of the parts in the interaction. What are the costs of these assumptions? Analytic narratives emphasize the importance of the institutional context in which situations take place. This institutional context can, in fact, provide the analytic narrativist with what Greif calls the "socially articulated and disseminated rules" that provide individuals with the cognitive, coordinative, and informational abilities that the theory assumes. Thus, the specificity of the institutional context can in many cases justify the behavioral assumptions (Greif, 2006, chap. 5).

The fact, however, that game theory may not be useful in certain conditions does not imply that analytic narratives must be abandoned. They can theorize by means of other forms of game theory that rely on different behavioral assumptions—for instance, evolutionary game theory and behavioral game theory. Uncertainty about the possible alternatives and complexity of the situation can lead to biases from the predicted game theoretic behavior. In these situations, the scholar needs to evaluate whether the theory should be modified to incorporate uncertainty or the pertinent behavioral assumptions by making use of behavioral or evolutionary game theoretic models. The usefulness of analytic narratives requires an awareness of the types of situations in which game theory is best applicable and that can benefit most from the methodology.

Analytic narratives, therefore, recognize and take advantage of the context specificity and historical contingency of institutional analysis. At the same time, they rely on explicit formal theorizing that compels scholars to take a systematic approach and reveal the reasoning and assumptions behind their explanations. This empirical methodology allows deduction and induction to complement each other and to be complemented by a context-specific analysis.

Luz Marina Arias
University of California, San Diego
La Jolla, California, United States

See also Analytic Narratives: Applications; Case Studies; Game Theory; Rational Choice

Further Readings

Alexandrova, A. (2009). When analytic narratives explain. *Journal of the Philosophy of History*, *3*, 1–24.

Bates, R. H., Greif, A., Levi, M., Rosenthal, J.-L., & Weingast, B. (1998). *Analytic narratives*. Princeton, NJ: Princeton University Press.

Greif, A. (2006). *Institutions and the path to the modern economy: Lessons from medieval trade*. Cambridge, UK: Cambridge University Press.

Kiser, E., & Welser, H. T. (2007). The microfoundations of analytic narratives. *Sociologica, 3*. Retrieved October 13, 2010, from http://www.sociologica.mulino.it/journal/article/index/Article/Journal:ARTICLE:119

Levi, M. (2003). An analytic narrative approach to puzzles and problems. In I. Shapiro, R. Smith, & T. Masoud (Eds.), *Problems and methods in the study of politics* (pp. 201–226). New York: Cambridge University Press.

ANARCHISM

Etymologically, the term *anarchism* derives from two ancient Greek words: αν, meaning "absence of," and Αρχή, meaning "authority," "government," "ruler," or "war chief." Anarchism thus carries a negative charge: the negation of authority, the absence of rulers. *Anarchism* is opposed to all forms of hierarchy, including the state, capitalism, religious institutions, patriarchy, and racism. Yet the term also conveys a positive political project of justice, liberty, equality, and solidarity. According to the basic principles of anarchism, genuine freedom, equality, and solidarity are logical and political impossibilities in a system or a regime where some rule and others are ruled. All political regimes except anarchy are oppressive and tyrannical, even those claiming to embody the nation or the sovereign people. That is why anarchism can be seen as the ultimate revolutionary ideology and social movement. From the perspective of all elitist ideologies, including liberalism, anarchism is fundamentally flawed because, the argument goes, human beings cannot find satisfaction unless they are under the authority of rulers, be they warlords, priests, nobles, property owners, or elected representatives.

As a political ideology or philosophy, anarchism involves three elements: (1) a positive ideal (anarchy), (2) a critical discourse (opposition to hierarchies and oppression), and (3) the means to achieve this ideal (evolution or revolution, nonviolence or violence, mass movement, or affinity groups). In some cases, it may also involve an ontological conception of human nature. By and large, anarchism adopts a neutral sociological outlook posited on a binary human nature with a positive and a negative pole. The positive pole is the ability to live with others in a spirit of equality and solidarity, while the negative pole is the selfish drive for domination. As demonstrated by Peter Kropotkin and Élisée Reclus, it is the social structure that determines which pole prevails. If the structure is hierarchical, then people in a position of authority will become self-indulgent and arrogant and abuse their authority. Only a social environment where no one is in a position to exert power over others can ensure that people's goodness will prevail.

This entry describes the evolution of anarchism from its early roots to its development during the "anarchist century" from the mid-19th century to 1939. It also examines the resurgence of anarchism in the latter part of the 20th century, influenced by postmaterialist thought.

Historical Development of Anarchism

Some commentators assert that anarchism as a political ideology or philosophy has existed since the dawn of humanity and can be found, at least to a degree, in Taoism, stoicism, the works of Zeno, and the concrete experiences of communities on every continent that practiced egalitarianism (among adult males, at any rate). Kropotkin's book *Mutual Aid* is instructive in this connection. Others consider anarchism as belonging essentially to European modernity, with its forerunners emerging in religious social uprisings (the Anabaptists), the English Revolution (the diggers), and the French Revolution (the *enragés*). William Godwin (1756–1836) is generally regarded as the first true anarchist philosopher (*Enquiry Concerning Political Justice*, 1793). The Frenchmen Anselme Bellegarrigue, editor of *L'anarchie—Journal de L'ordre* (of which only two issues appeared, in 1850), and Pierre-Joseph Proudhon were among the first to call themselves "anarchists." They were followed by Mikhail Bakunin (1814–1876), Louise Michel (1830–1905), Voltairine de Cleyre (1866–1912), Peter Kropotkin (1842–1921), Élisée Reclus (1830–1905), Emma Goldman (1869–1940), and Errico Malatesta (1853–1932).

In its early days, anarchism attracted several prestigious fellow travelers, such as the painters Gustave Courbet and Camille Pissarro, the playwright Henrik Ibsen, the literary critic Herbert Read, and the philosophers Emmanuel Mounier and Bertrand Russell. Some famous anarchists, like the geographers Kropotkin and Reclus, were eminent scientists. Contemporary scholars identified with anarchism include Allan Antliff (art history), Normand Baillargeon (education science), Harold Barclay (anthropology), Susan Brown (political philosophy), John Clark (philosophy), David Colson (sociology), Ronald Creagh (sociology), Uri Gordon (social ecology), David Graeber (anthropology), John Holloway (political philosophy), Ruth Kinna (political science), and Robert Wolff (political philosophy).

There is an anarchist studies network linked with the Political Science Association in Britain and a number of scholarly anarchist journals, such as *Anarchist Studies*, *Réfractions*, and *Social Anarchism*. Yet it is in the nature of anarchism to expect intellectual and political activities to go hand in hand (*praxism*), and any division of labor between theorists or scholars and activists is viewed with suspicion. Bakunin held that anarchism acknowledges that some people may be more talented or skilled in a specific area, including theorization. However, there must be no monopoly of expertise. To be free, the people dealing with any expert should have the following: first, access to other experts in the same field; second, the right to ignore the advice of experts; and, finally, the opportunity to become experts themselves through education and practice. In other words, while expertise may confer influence on its possessor, it should be available to all and not be used to wield power over others. Anarchist ideas are for the most part expressed and disseminated anonymously in self-published independent journals, zines, and websites (see Infoshop) or in the publications of anarchist groups and organizations (e.g., Fédération Anarchiste and Alternative Libertaire in France, Anarchist Black Cross and Northeastern Federation of Anarchist-Communists in the United States, Class War in Great Britain, and Union Communiste Libertaire au Québec in Canada). Individual artists, like the composer John Cage, or trends, such as the punk movement (e.g., Bérurier Noir, Crass), are also identified with anarchism.

Central Elements of Anarchism

The Classical Age of Anarchism

From a historical perspective, there are two distinct moments of anarchism as a social and political movement. First, the classical age of anarchism—or the "anarchist century"—began with the first texts of Proudhon (who in 1840 famously stated, "Property is theft") and ended with the defeat of the Spanish Revolution in 1939. The period was characterized primarily by the working-class's expression and experience of anarchism, although some activists were already concerned with issues such as women's emancipation (abortion, birth control), free love and bi- or homosexuality, war and peace, racism and anticolonialism, free education, and vegetarianism. This epoch encompasses the Paris Commune of 1871, the Industrial Workers of the World in the United States, the Makhnovist rebellion during the Russian civil war, and the Spanish Revolution (1936–1939). During this time, anarchism was influential in Central and Latin America, especially in Argentina from the late 1800s through the 1920s. Of note is the year 1919, particularly the "bloody week," when about 1,000 people were killed and 50,000 arrested during a strike, including many anarchists. In Chile, anarchist organizations boasted 50,000 members in 1910 out of a total population of 3 million. In Mexico, anarchism was known through popular figures such as Ricardo Flores Magon (1874–1922), while in Cuba there were tens of thousands of activists in the 1920s. Anarchist activists and propagandists were very often at the forefront of social struggles and strike movements, such as the campaign for the 8-hour working day. In other parts of the world, anarchism was limited to small groups, active only for short periods of time but nevertheless subjected on occasion to bloody repression. Examples include the Philippines anticolonialist movement around 1900 as well as the Japanese anarchist and antiwar activist Osugi Sakae (1885–1923) and the anarcha-feminist Itô Noe (1895–1923).

The classical age of anarchism is still widely associated with illegalism (social banditism) and terrorism ("propaganda by deeds"). Around 1900, terrorists assassinated several heads of state (French President Sadi Carnot, 1897; William McKinley, president of the United States, 1901; Russian Prime Minister Pyotr Stolypin, 1911; Spanish Prime Minister José Canalejas, 1919; King George I of Greece, 1913), property owners, and military officers, very often in response to the bloody repression of the workers' movement. But at no time did so-called anarchist violence even come close to the level of violence (imprisonment, torture, death penalty, mass murders, etc.) perpetrated against anarchists by agents of the state, private paramilitary units and death squads, nationalist and fascist militias, and authoritarian Marxist militants. Anarchists were targeted by special repressive laws in almost all Western states in the closing decade of the 19th century. Thousands were killed by their former allies in Russia (Leninists) and Spain (Stalinists), and they were among the first political prisoners to enter the Nazi concentration camps in Germany. Violent repression, in conjunction with the heightened popularity and power of authoritarian Marxists following the 1917 Russian Revolution and the Red Army resistance and victory against the Nazi armies in 1941 to 1945, accounts for anarchism's declining influence within the progressive and revolutionary movements during the first half of the 20th century.

Postmaterialism and the Revival of Anarchism

The second moment of anarchism relates to the so-called postmaterialist context of the 1960s.

With spectacular events such as May 1968 in Paris, anarchism experienced a revival, although authoritarian Marxists (Leninists, Maoists, Trotskists, Guevarists, etc.) remained very influential among radicals. The anarchist resurgence was also connected with environmentalism, feminism, and pacifism, which embodied anarchism not only by advocating the abolition of capitalism, patriarchy, and the state but also by organizing their militant activities with neither formal leaders nor hierarchy, through a strictly egalitarian, participative, and deliberative decision-making process.

The contemporary resurgence of anarchism as a major tendency among radical progressive forces came with the collapse of the Soviet empire, the intellectual crisis of Marxism, and a series of events at the turn of the 21st century that signaled the rise of the so-called antiglobalization movement and resulted in part from anarchist mobilization: the World Social Forum in Porto Alegre, Brazil (2001) and the mass protests in Seattle, Washington, in 1999; Washington, D.C., and Prague, Czech Republic, in 2000; and Québec, Canada, and Genoa, Italy, in 2001. Many commentators saw this as a "new anarchism," characterized by (a) the innovative tactics of direct action, (b) the horizontal structure of militant organizations, and (c) the inclusion of pluralist concerns. The direct action tactics were exemplified by the Black Blocs, a street tactic originating in the German *Autonomen* movement of the 1980s; Reclaim the Streets in the United Kingdom; and carnivalesque actions such as the Pink Blocs or the Clandestine Insurgent Rebel Clown Army. The horizontal structures were inherited from the previous generation of radical ecologists, feminists, and pacifists. Finally, the pluralist concerns, emphasizing cultural and sexual diversity, were clearly due to the influence of the feminist and gay and lesbian movements, although some anarchists were still based in the labor movement or were actively involved in labor issues and anticapitalist mobilizations. In principle, the neo-anarchists (or "postanarchists"—Hakim Bey, Lewis Call, Todd May, and Saul Newman) regard as equally important the need to oppose the state, war, capitalism and neoliberalism, racism, sexism and heteronormativity, ageism, and specism. This opposition must furthermore take place globally as well as locally and in everyday life. Indigenous movements of resistance, such as the 1994 Zapatista insurrection in Chiapas, Mexico, are viewed as stimulating examples of anarchism in practice, mainly because they are self-organized through participative and deliberative popular assemblies (the Mohawk scholar Taiaiake Alfred has coined the term *anarcho-indigenism*).

Currents in Anarchist Thought

Historically, there have been several anarchist currents, each with its particular position on what is primary (the individual or the community, liberty or equality), its priorities (the emancipation of workers or women, the environment), and its organizational structure (mass movements, unions or communes, affinity groups). Very often, they have collaborated in solidarity, at times they have engaged in intense debates, and sometimes they have clashed.

Individual(ist) Anarchism

Individualist anarchism, according to Max Stirner, 1806–1856; Lysander Spooner, 1808–1887; and Benjamin Tucker, 1854–1939, emphasizes liberty and autonomy and denies that there is something other than individuals to which an individual must submit (e.g., nation/motherland, class/proletariat). However, such anarchists are as a rule deeply concerned by the exploitation of the working class and the oppression of women and may very well take part in social struggles, but through affinity groups rather than mass movements. They generally believe that free association can be beneficial for the individual's happiness and interests, and through the years they have been involved in rural or urban communes and in squats. However, membership in an association must be totally voluntary, with the freedom to join or leave at any time.

Libertarianism and *anarcho-capitalism* are sometimes used as synonyms of individual anarchism, although the latter is in fact a kind of radical liberalism, advocating free market capitalism for property owners and wageworkers, under the protection of either a state whose function is strictly limited to the enforcement of law and order or, in the absence of a state, of private security agencies. These views imply that property owners are in a position of authority and power over their employees, an idea inconsistent with individual as well as collective anarchism. *Proprietarian* is the term coined by Murray Bookchin to describe libertarianism.

Like Marxism, anarchism views capitalism as an exploitative system, protected by the state, which also exploits workers to sustain itself. This said, anarchism offers a variety of economic proposals for abolishing capitalism and reorganizing the production of goods and services.

Collectivist Anarchism

Proudhon proposed mutualism (from the Mutualists, a 19th-century secret society of weavers in Lyon, France). He suggested that just as the hive belongs to all bees, the means of production and exchange must be collective. However, each member of the collective—a free association—should share to some degree in the worst tasks. This job rotation would provide everyone with a better knowledge and understanding of the whole working process. Higher functionaries, such as architects and engineers, would be assigned through election, but decisions and working rules would be made collectively through a participative deliberative process. The exchange value would be assessed and then converted into tickets of hours of work under the supervision of a People's Bank, which would also offer interest-free credit. Consequently, there is no need for a violent revolution; workers need only organize among themselves to be free.

Some anarchists have identified serious issues arising from Proudhon's economic proposal: Should the evaluation of hours of works take into account job intensity and risk, required training and skills, and the worker's social responsibilities (such as dependents)? Moreover, this remains a very individualistic model of work organization, which draws on a bygone tradition of craftsmanship that values the artisan's self-reliance and personal pride. Finally, according to some, Proudhon is not a true anarchist, because he devotes considerable energy to demonstrating the physical, intellectual, moral, and political superiority of men over women, as evidenced by the more than 500 pages of blatant misogyny and antifeminism in his *De la Justice dans la Révolution et l'Église* (1860).

Anarcho-Syndicalist Anarchism

Anarcho-syndicalists, identified with Rudolf Rocker (1873–1958), see labor unions (*syndicats* in French) as the spearhead of social emancipation and revolution and the seed of the future anarchist society. Property owners—the bourgeoisie—and the state create nothing; they are mere parasites of workers' productive activity. Hence, if workers unite and stop working for their exploiters, the system will collapse and revolution will ensue. The general strike is therefore the main instrument for the realization of anarchy. Thereafter, global production will be structured according to "industrial groups" in each sector of production and integrated in an "economic federation." However, collective decision making would remain a bottom-up process centered on the general assembly of every place of work. Economic production based on human needs will ensure that all receive the goods and services they require.

Anarcho-Communism

With anarcho-communism, identified with Peter Kropotkin and Alexander Berkman (1870–1936), among others, the state becomes the direct target of a revolutionary mass movement. Although anarcho-communism often stresses the importance of the workers' movement, it is more open than many other anarchist trends to peasants, women, and the *lumpenproletariat*. Adopting a geographical perspective, anarcho-communism regards the urban or rural commune as the center of the new society. After the revolution, the communes may join in a federation, although the decision-making function will be situated at the local level, in keeping with a bottom-up process. In the absence of exploiters, such as the bourgeoisie and state agents, production output will necessarily surpass society's needs. Anarchy will therefore lead to a situation of abundance where all will receive whatever goods and services they need and want.

Working Inside Versus Outside the Political System

"Libertarian municipalism" is a contemporary version of anarchism, primarily developed in the 1980s by Janet Biehl and Murray Bookchin. Libertarian municipalism lays special emphasis on ecological issues, contending that these could be dealt with more effectively at the local level but only if municipal institutions were deliberative and participative. Bookchin even suggests that anarchists get involved in municipal politics, a view that has been criticized by other anarchists as obviously inconsistent with the antistate anarchist principle.

Bookchin is not the only self-proclaimed anarchist calling for some sort of involvement in official politics. Authors and activists like Colin Ward and Paul Goodman, or Philippe Corcuff, in France, who coined the term *libertarian social democracy*, argue that some anarchism in the official system is better than no anarchism at all. They advocate reforms with regard to issues, laws and regulations concerning individual freedom, free education, free sexuality, and so on. According to Noam Chomsky, it is the duty of today's anarchists to support the state and its welfare programs against neoliberalism and capitalism, or what he calls "private tyrannies." Chomsky is even in favor of voting for the Democratic Party in the United States, though he admits that mass social movements are still necessary to push the political elite in the right direction.

Others argue that it is possible to have anarchy here and now but outside the official system. For instance, the green anarchist Clark advocates "microcommunities." In today's network of squats, for instance, activists who align themselves with anarchism take over unoccupied buildings or land and try to live outside capitalism as much as possible, offering free meals and rooms to travelers and traveling to participate in campaigns of solidarity with immigrants or with antiwar and anticapitalist movements.

Feminism and Anarchism

The distinctions between various forms of anarchism are not as schematic in reality as what is generally presented in books, articles, and encyclopedias. Other strands of anarchism do not fit perfectly within the classical economic and political framework. Anarcha-feminism—for instance, is a feminist version of anarchism, devoted to fighting patriarchy both in society at large and within the Marxist and anarchist networks, where misogynistic and antifeminist attitudes are not uncommon (witness the term *manarchy*, recently coined by U.S. activists to describe inequalities among anarchist men and women). In the 19th century, authors and militants like Goldman and de Cleyre denounced the "sex slavery" of marriage and promoted free love. In the 1930s, the anarcha-feminist autonomous association *Mujeres Libres* boasted some 30,000 members in Spain and was an active force during the Revolution. It focused on education for girls and women, combat training for women, medical help for wounded fighters, and the emancipation of prostitutes through psychological and material support and the assassination of certain pimps. From the 1960s onward, Peggy Kornegger and Susan Brown, among others, highlighted the similarities between anarchism and feminism, arguing that, to be consistent, anarchists should be feminists (although some feminists, such as the liberal or Marxist statist feminists, may not be anarchists, i.e., opposing all forms of hierarchy and inequality). "Third-wave" feminism, with its emphasis on individual sexual choice and *queer* identities, is sometimes identified with anarchism. Nevertheless, the anarcha-feminist Claire Snyder warns "choice feminists" that under patriarchy, women's sexual life may not be totally free from the external control of men or from "internal tyrants," the name given by Goldman to the socialization—internal fears and false beliefs—that undermines women's freedom of will and choice. Finally, with regard to economic production, the anarchist and feminist Carol Ehrlich notes that it is not only the bourgeoisie as a class that exploits wageworkers but also men who exploit women's work, which is often undertaken with no monetary compensation.

Anarchism in the 21st Century

Today, a large number of principles closely linked to anarchism for the past 150 years have become, for many people, ordinary moral and political liberalism: the 8-hour working day; education for all, boys and girls alike; women's freedom of choice with regard to matters such as contraception and abortion; freedom in sexuality and love, including bi- and homosexual rights and the right to divorce or to cohabit without marrying; freedom of consciousness with respect to religion and state dogma; and freedom to refuse military service. While not self-identifying with anarchism, several influential contemporary scholars, especially in France and Italy, have adopted an anarchist approach in their work on state power (Giorgi Agamben), political representation (Michel Foucault, Jean-François Lyotard), micropolitics and resistance (Gilles Deleuze), and the reorganization of production and resistance based on egalitarian principles (Michael Hardt and Antonio Negri). Others, such as Donatella della Porta, are investigating new social movements that practice "direct democracy," which in fact amounts to studying anarchism in action.

This implicitly or explicitly anarchist activism within the "movement for global justice" is the target of state repression, even though anarchism does not represent a significant threat to social, political, and economic liberalism, as might have been the case at the turn of the 20th century. Thus, there were legitimate grounds for former British Prime Minister Tony Blair's jibe about the summit-hopping "anarchist traveling circus." Nevertheless, anarchism provides the most vocal and radical left-wing criticism of liberalism. During the 2001 G8 (Group of Eight) Summit in Genoa, after violent clashes between police officers and protesters—one of whom was shot dead at point-blank range—Canadian Prime Minister Jean Chrétien warned about the "anarchists who want to destroy democracy." In Greece and Italy, anarchist networks have been targeted by the police in recent years: Many have been arrested, and in Greece, one person was killed in 2008. The Federal Bureau of Investigation and Scotland Yard have also been preoccupied since the late 1990s with anarchists reputedly training for terrorist actions (which never occurred). Even certain political scientists, such as Tim Dunne of Exeter University, claim that anarchist activists and Al Qaeda militants are very similar in nature. In sum, anarchism said goodbye to the 20th century much as it had welcomed it—that is, cast in the role of a potential "terrorist" threat, the number one internal public enemy.

Francis Dupuis-Déri
Université du Québec à Montréal
Montreal, Quebec, Canada

See also Anarchy; Capitalism; Democracy, Direct; Libertarianism; Power; State; Terrorism, International

Further Readings

Gay, K., & Gay, M. K. (Eds.). (1999). *Encyclopedia of political anarchy*. Santa Barbara, CA: ABC-CLIO.
Gordon, U. (2007). *Anarchy alive! Anti-authoritarian politics from practice to theory*. London: Pluto Press.
Graeber, D. (2004). *Fragments of an anarchist anthropology*. Chicago: Prickly Paradigm Press.
Jun, N. J., & Wahl, S. (Eds.). (2009). *New perspectives on anarchism*. Lanham, MD: Lexington Books.
Marshall, P. (2007). *Demanding the impossible: A history of anarchism*. Glasgow, UK: HarperPerennial.
Miller, D. (1984). *Anarchism*. London: Dent.
Pennock, J. R., & Chapman, J. W. (Eds.). (1978). *Anarchism* (*Nomos*, No. 19). New York: New York University Press.
Rolling Thunder journal: http://crimethinc.com/rt
Social Anarchism journal: http://www.socialanarchism.org
Star, D. (Ed.). (2002). *Quiet rumours: An anarcha-feminist reader*. Edinburg, UK: AK Press.

Anarcho-Syndicalism

See Anarchism; Anarchy

Anarchy

Anarchy is the absence of government or, more generally, political authority over and between the units of a political system. As an analytic concept, the term does not imply a lack of political order or the presence of chaos and thus differs from informal and colloquial use. The term also differs from *anarchism*, a normative and possibly utopian position that advocates minimizing the scope of political authority to maximize the domain of individual autonomy.

The condition of anarchy is widely understood to describe the modern international system in which states are the units of analysis, each is fully sovereign, and all are formally equal. It is this condition of anarchy, in turn, that separates international relations from other domains of politics and renders it, for many analysts, a distinct field of inquiry with different rules and patterns of interaction. Although other political arenas may also be anarchic, such as legislatures where vote trading between members cannot be legally enforced, the analytic concept has not been widely applied beyond the realm of international relations and, in a few cases, failed states.

Consequences of Anarchy

For most scholars of international relations, anarchy requires that all states must rely only on their own resources and abilities—a practice described as the principle of self-help. Lacking any authority that states can appeal to for protection, aid, or binding adjudication, each state must ultimately depend on its own efforts and wisdom. This further implies that any agreement made between states must be self-enforcing or in the interests of the parties to carry out if and when actually called on to do so.

Although the assumption that the international system is anarchic is widely shared, the meaning and consequences of this assumption are still vigorously debated. For realists, anarchy produces a zero-sum, competitive struggle between states. For political realists, anarchy did not feature prominently and was more a passive background condition; Hans Morgenthau—for instance, rooted the drive for power in the innate character of political man, not in the nature of the international system. For Neorealists (sometimes called structural realists), however, anarchy is one of the defining features of international structure with significant causal effects.

Even within neorealism, there are two prominent schools. For defensive realists, anarchy requires only that states seek security, although they may also pursue expansion, glory, or power for other reasons. Given that some states may possess aggressive tendencies, however, all states must be ever vigilant and prepared to defend themselves. Uncertainty over the motives of other states and problems of credible commitment, in turn, sometimes leads to bargaining failures and war. Offensive realists believe anarchy is such a challenging condition that states must pursue power at all times. Since states are always insecure, the fear that others will exploit them forces states to pursue all means necessary to impose their will on others. Also, as power is always zero sum, anything that gives an advantage to one state must create a disadvantage for at least one other. In this view, anarchy implies that international politics are a perpetual and intense struggle for domination.

Neoliberal institutionalists see anarchy as a condition that can be mitigated, if not fully resolved, by voluntarily negotiated institutions between states. Unlike in neorealism, anarchy does not define the goals that states seek but, rather, merely permits dilemmas of collaboration and coordination to arise that thwart cooperation between utility-maximizing states. These dilemmas, in turn, can be overcome in part by institutions that provide information or make commitments more credible. In a form of self-organizing order, states can avoid some of the harsher implications of anarchy by building institutions without necessarily subordinating themselves to any central authority.

Finally, constructivists see anarchy as an open-ended condition filled by the social purposes of states, which are themselves constructed of socially appropriate and interpreted roles and norms that vary over time and space. Thus, within the condition of anarchy, states may represent themselves and others in ways that produce a Hobbesian world depicted by realism, a Lockean world that reflects important elements of neoliberal institutionalism in its focus on natural rights, or a Kantian world of greater peace and cooperation than imagined by either of the alternatives. In all of these different socially constructed world orders, there is a potential for systemic transformation that is excluded by neorealism or neoliberal institutionalism.

Critiques of Anarchy

Although subject to continuing controversy, all three perspectives share a common focus on anarchy even while they debate its meaning and implications. Increasingly, scholars are criticizing the very concept of anarchy and the role it has played in the development of international relations theory. Too numerous to describe in detail here, many of the specific critiques can be grouped into two larger themes. First, the common notion of anarchy is derived from an overly narrow formal-legal conception of authority. Alternative conceptions of authority open up the possibility of many different kinds of authority existing at the same time within the international system. Second, sovereignty is not indivisible, as commonly averred, but is a bundle of different authorities that can be disaggregated in ways that do not necessarily coincide with traditional nation-states. Divisible sovereignty permits a patchwork of sometimes overlapping, competing, or complementary authorities to exist simultaneously. Taken together, these critical themes suggest that many forms of global governance that have been ruled out by the common assumption of anarchy may have actually existed in the past and may yet expand in the future.

In formal-legal conceptions of authority, the person (or unit) in authority has the right to issue and enforce certain commands over a set of subordinates because of the lawful position or office that he or she holds. Authority does not inhere in the individual (or individual unit) but in the person as an officer who is duly appointed or elected through some lawful procedure. Elected by a majority of the Electoral College, for instance, a person becomes the president of the United States; by established rules of rotation, a country collectively holds the presidency of the Council of the European Union (EU). Applied to international relations, since there is no lawful authority or procedure above states, no state or other unit can be authorized to govern over other states. As a result, the system and relations between states within that system must be anarchic.

Formal-legal authority, however, is simply one possible source of authority. The German sociologist Max Weber, closely identified with this formal-legal conception, also posited that legitimacy and authority could derive from charisma, tradition, or religious belief as well. Others argue that legitimacy and, hence, authority can arise from psychological principles of fairness and justice, socially constructed norms, or a negotiated social contact between ruler and ruled. These other sources of legitimacy at least open the possibility of authority by states, international organizations, nongovernmental organizations (NGOs), and other units over other similar units. While debate continues over who has authority over whom for what, a growing body of critical research suggests that international relations are not entirely anarchic but better described as a variegated system of multiple units in authority drawn from multiple sources of authority.

In turn, sovereignty has been traditionally conceived as indivisible or culminating in a single apex within each territorially distinct entity. Sovereignty was once vested in a king, emperor, or "sovereign." Today, sovereignty typically resides in the "people" and has itself evolved over time to include most citizens of a country. But in either case, there is assumed to be within each society a single, ultimate authority that cannot, even in part, be subordinate to any other actor. This assumption is codified in the notion of Westphalian or juridical sovereignty embodied in the United Nations Charter; today, states need not actually control their territory, as in the past, but only need to be recognized as sovereign by other sovereigns to be accorded that status.

Yet there is an increasing awareness that states have never fit this idealized vision of Westphalian sovereignty. Indeed, even historically, sovereignty appears better described as a set of authorities that are disaggregated across units within a state, long recognized as federalism, or, more important, across states and third parties. This disaggregation is not only exemplified by multilevel governance in the contemporary EU but also reflected in numerous international restraints on the freedom of action of states, including rights of protection or guarantee (e.g., the United States and the Federated States of Micronesia), rights of economic and financial control (the United States and the Dominican Republic from 1904 to 1941), rights of servitude (the United States and Japan, under various status of forces agreements), and rights of intervention (the United States and Panama, under the neutrality treaty of 1977).

Combined with multiple forms of authority, the possibility of multiple sites of authority suggest that patterns of global governance are likely to be more varied, complicated, and dynamic than once assumed. Transnational nonstate groups exercise authority over their members, whether they be religious orders, labor unions, or other collective bodies. NGOs earn authority over firms and even states in standard-setting boards and international creditor cartels and regulate behavior through monitoring and certification procedures. States exert authority over one another in spheres of influence, protectorates, or informal empires. International organizations possess authority as well, including the World Trade Organization and its dispute settlement procedures over trade, for instance, and the United Nations and the International Atomic Energy Agency over nuclear facilities and programs.

The assumption of anarchy, critics charge, blinds scholars of international relations to this variegated system of global governance. The critique of anarchy and the expansion of global governance implies three essential challenges for the future: (1) Analysts must map the forms of global governance, identify where current authority is inadequate or itself unregulated (as in the so-called democratic deficit in the EU), and propose reforms to improve human welfare; (2) policymakers must accept and navigate between these multiple forms of authority while harnessing them to their national purposes, when appropriate; and (3) global citizens must work to ensure that global authorities act in the general interest.

David A. Lake
University of California, San Diego
La Jolla, California, United States

See also Anarchism; Constructivism in International Relations; Governance, Global; Governance, Multilevel; International Organizations; Legitimacy; Neoliberal Institutionalism; Realism in International Relations; Sovereignty; State Failure; Weber, Max; Westphalian Ideal State

Further Readings

Keohane, R. O. (1984). *After hegemony: Cooperation and discord in the world political economy.* Princeton, NJ: Princeton University Press.
Krasner, S. D. (1999). *Sovereignty: Organized hypocrisy.* Princeton, NJ: Princeton University Press.
Lake, D. A. (2009). *Hierarchy in international relations.* Ithaca, NY: Cornell University Press.
Milner, H. V. (1991). The assumption of anarchy in international relations: A critique. *Review of International Studies, 17*, 67–85.
Waltz, K. N. (1979). *Theory of international politics.* Reading, MA: Addison-Wesley.
Wendt, A. (1999). *Social theory of international politics.* New York: Cambridge University Press.

Anomia

Anomia is a noun that comes from the Greek *a* (negative prefix) and *nomos* (law), meaning absence of norms. The original meaning of anomia in the frame of the classical Greek polity defined a condition of being against or outside the law or a situation where current laws were not applied, resulting in a state of illegitimacy or lawlessness. In its modern usage, the meaning of anomia has a much broader scope, although it is a rather recent concept in the history of social thought. In much of the literature, particularly in the English-speaking world, the French spelling *anomie* is also extensively used. This entry discusses the use of the term in modern sociology, the widening of its scope, and its present-day relevance.

Anomia in Durkheim's Thought

In the social sciences, anomia has been a key concept in the development of modern empirical sociology. It is normally associated with the work of Émile Durkheim (1858–1919), the French sociologist who introduced the concept in his early sociological masterpieces, *The Division of Labor* (1893) and *Suicide* (1897). In his sociological perspective, anomia was no longer related to the rule of law; thus, it lost its original legal content and was instead defined as the absence of norms (legal rules just being one class of social norms). One of the major tenets of Durkheimian sociology is that a theoretical explanation of social agency must be given in terms of social causal factors and not on the basis of psychological conditions. In fact, one can say that the major contribution of Durkheim to the foundation of modern sociology has been the possibility of grounding the explanation of social action on purely social variables. In the famous Durkheimian formula, presented in Chapter 5 of *Rules of Sociological Method* (1895), social facts must be explained by other antecedent social facts. In relation to anomia, this thesis implies that the absence of norms is not a subjective psychological condition of the agent but an objective feature of the social structure or, in strictly Durkheimian terms, of the *collective life*.

The first references to anomia by Durkheim appear in *The Division of Labor*, his doctoral thesis at the University of Bordeaux. In this early work, referring to the social causes of the division of labor as a phenomenon typical of modern industrial society, Durkheim elaborates on the "abnormal" forms of this phenomenon, one of these forms being the "anomic division of labor."

In his ensuing famous essay *Suicide*, in which Durkheim put forward a plausible sociological explanation of the apparently purely psychological phenomenon of disposing of one's own life, he characterized a type of suicide that he called "anomic." This type occurs in the case of major social crises (e.g., an economic crash)—that is, structural phenomena that invalidate previous regulations to such a degree that some agents are incapable of coping with the radical ambiguity of the ensuing state of normlessness, and thus they take their own life in a supreme act of "anomic suicide." The common factor of both the anomic division of labor and the anomic suicide is the dissolution of social integration, the underlying social phenomenon that is manifested in the absence of norms. The reformist stance of Durkheim vis-à-vis the major political and social issues of his time is a reflection on how it were possible to confront trends toward anomic forms of social life in a secularized and functionally differentiated modern society.

In one of his most politically oriented works, *Leçons de Sociologie* (1950), published in English as *Professional Ethics and Civic Morals* (1957), Durkheim analyzes the relationships between an individual and the sphere of public regulation. In his view, the institutionalization of individual rights is the role of the State; thus, the latter has not been created to guarantee that individuals are able to exercise their natural rights; rather, the State creates and organizes those fundamental rights, and it makes them a reality. Political anomia is then a condition where public regulations have reached a point of extreme weakness, close to the absence of rules, and, therefore, individual rights cannot be exercised. In political terms, however, anomia is not anarchy. The latter refers to a lack of leadership, hierarchy, and command, whereas anomia implies a weakening of rules, structures, and organizations. Clearly, the Durkheimian concept of political anomia necessarily implies a radical weakening of public regulations. In short, anomia, in all of its possible specific manifestations, can be understood as a societal consequence of rapid and deep social change, and it is in this sense that anomia is a concept that runs in parallel with the notion of social crisis. According to Durkheim, without clear rules individuals cannot find their place in society. Such a state tends to occur in periods of social disruption (such as rapid economic growth or economic depression), bringing about greater anomia and, concurrently, higher rates of suicide, crime, and deviance.

Anomia and Social Change

Social change may generate anomia either in the whole of society or in some parts of it, so it must be stressed that, in a given society, social groups may be differentially affected by anomia. As recent events in the world economy illustrate, economic crises brought about by business cycles have differential impacts on groups according to their location in the social pyramid. Both sudden downward mobility, which is normally associated with economic depression, and the rapid onset of material prosperity, leading to quick upward mobility, tend to upset previous networks in which lifestyles are embedded, in each case increasing the chance of anomia. Talcott Parsons, in the early 1940s, discussing the social aspects of fascism, contends that an increase in anomia may be an outcome of almost any social changes that upset previous routines of life or symbolic associations, all of which imply instability of expectations. In Parsons's strain theory approach, social ambiguity as well as conflicting role expectations may lead to increased anomia, and this, in turn, may create a sort of anxiety-generated pattern of anomic action in areas such as reactionary voting behavior. In expanding the use of the concept of anomia to diverse areas of social interaction, Edward Tiryakian has made an elaboration of the notion of sexual anomia as an heuristic tool in approaching features of contemporary society that involve transformations in sexual relations and sexual identity that presuppose deregulations in modern sexual relations.

Anomia and Deviant Behavior in Merton's Theory

A distinction in the normative sphere between regulation of goals and regulation of means, introduced by Robert K. Merton, has become a key dimension in current anomia-related theories that tend to focus on "anomic" deviant behavior. A clear example of this shift in focus is the way in which the *Columbia Encyclopedia* (6th ed.) defines anomia: "a social condition characterized by instability, the breakdown of social norms, institutional disorganization, and a divorce between socially valid goals and available means for achieving them." Although Merton, when referring to socially accepted goals (values) and means (instruments for achieving those values), has in mind the overall universal striving for success in the American value system, his theory is a general proposition. In Mertonian terms, in any given society, there is a range of culturally accepted values (goals) and a corresponding set of normatively approved means of securing these goals, but the structure of economic resources in that society enables only certain privileged groups and classes to succeed.

Disparity between goals and means is, according to Merton, the main cause of anomia. Merton identified five possible reactions to the different relationships between goals and means. The first and most common reaction is *conformity*, which presupposes a nonanomic relationship of appropriate means to reach the socially accepted goals. All of the other four possible reactions are more or less anomic. The second possible reaction is *innovation*, which implies adherence to goals and rejection of socially accepted means. The third possible reaction is r*itualism*, based on adherence to means but rejection of socially accepted goals. In the fourth reaction, *retreatism*, both the goals and the means are rejected. Finally, the fifth and final possible reaction is *rebellion,* when both goals and means are ambiguously dealt with. In conditions of anomia (maladjustment between goals and means), underprivileged, deprived individuals will turn to various forms of individual deviance that are dependent on alternative means to reach the same desired ends. In other words, anomia occurs as the disjunction of means and goals that differ according to the profile of the social opportunity structure. Merton was then able to propose a theory of anomia embedded in social stratification and the class structure, and it is in this sense that his theory transcends the specificity of the American case and becomes a general theory of anomia. An interesting example of a Merton-inspired research of anomic behavior in a non-Western cultural context is the study conducted in 2006 by Parviz Piran, concerning the reactions of Iranian youth to anomic conditions created by the rapid and extensive change in the economic, political, and cultural fabric of their society.

Anomia in Contemporary Political Economy

The concept of anomia is also present in the field of contemporary political economy in the work of analysts who are advancing a theory of "institutional anomia" that basically states that in the present globalized context, as economic institutions gain dominance in shaping society, noneconomic institutions are weakened and forced to accommodate the market, thus feeding a condition of high institutional anomia that is conducive to higher violent crime rates. To the extent that increased anomia tends to erode and dissolve social networks in which social action is embedded, there is also a clear connection of this concept with the notion of social capital that is basically the outcome of associative networking. Contemporary research on the sources and the uses of social capital by different social actors in promoting their own advancement, which is a strong field in policy-oriented studies, is an interesting area for further applications of the general theory of anomia.

In short, the concept of anomia has provided quite a fertile intellectual ground for the development of the modern disciplines of the social sciences.

Raul Atria
University of Chile
Santiago, Chile

See also Alienation; Anarchy; Apathy; Crisis; Durkheim, Émile; Social Capital

Further Readings

Daday, J., Broidy, L. M., & Willits, D. (2007, November). *Institutional-anomie, political corruption and homicide rates.* Paper presented at the annual meeting of the American Society of Criminology, Atlanta, GA.

Johns, D. (2007, November). *Expanding Durkheim: Toward a graduated anomie theory of white-collar crime.* Paper presented at the annual meeting of the American Society of Criminology, Atlanta, GA.

Merton, R. K. (1957). Social structure and anomie. In *Social theory and social structure* (pp. 185–214). New York: Free Press.

Orrù, M. (1897). *Anomie: History and meanings.* London: Unwin & Hyman.

Parsons, T. (1942). Some sociological aspects of fascist movements. *Social Forces, 21*, 138–147.
Parviz, P. (2006). *Iranian youth and social transformation: Review of a research.* Tehran, Iran: Allameh Tabatabaee University.
Tiryakian, E. A. (1981). Sexual anomie, social structure, societal change. *Social Forces, 59*(4), 1025–1053.

APATHY

Apathy looks at first like a very broad, even vague concept. It plays, however, a very important and critical role in political analysis, particularly in the analysis of the decline in voting turnout and civic engagement and more generally in the study of participation and civic involvement of citizens. The concept is derived from a long, intellectual, interdisciplinary tradition. It was first proposed in philosophy but has been extensively used in cultural psychology, anthropology, sociology, and political science. The concept also has intellectual connections with other concepts or notions coming from the sociological and political philosophical traditions such as alienation and anomie. In a typical use of the apathy concept, theories of mass society have explained, in the wake of World War II, that modernization and urbanization have been responsible for detaching individuals from their primary groups (community, family, early socialization groups), generating withdrawal from the public sphere and lack of interest and motivation in public affairs. This entry discusses the history and recent criticisms of this concept.

The more recent perspective on political apathy started in the 1950s and 1960s, in particular with the publication of the book *Civic Culture* by Gabriel Almond and Sidney Verba (1963). One of the issues of Almond's and Verba's political culture approach is that the stability of democracy and democratic institutions depends on a "civic culture," a mixture of participatory, subject, and parochial orientations toward politics. The important point here is that this civic culture seems to allow many citizens to be apathetic about politics. Many critics of the book took this to be a kind of justification for apathy.

Since the 1960s, the political sociology perspective on apathy is mainly linked to the question of political mobilization and political participation. In a very challenging study of the relationship between social class and voting in the mid-1980s in the Netherlands, Leo B. Van Snippenburg and Peer Scheepers recalled that a positive correlation between social protest and income inequalities at least implicitly referred to the thesis of Verelendung (pauperization) that originated from Karl Marx's work. They also recall that this thesis states that under (relatively) deteriorating socioeconomic conditions, the deprived will turn to social protest and revolt. The thesis was later criticized, emphasizing that history has regularly shown the contrary: Revolt emerges under circumstances of relative social deprivation and is more likely to happen in middle-class or privileged groups rather than in deprived ones. According to these alternative views, instead of revolting against the system that generates inequalities, a form of "collective apathy" often manifests itself in deprived groups (Marie Jahoda, Paul Lazarsfeld, & Hans Zeisel, 1971). This doubt and criticism of the pauperization thesis was strongly formulated in studies by a very influential group of social scientists in the 1930s and the 1940s, the so-called Frankfurt School, in particular their study of the "authoritarian personality" (Theodor Adorno, Else Frenkel-Brunswik, Daniel Levinson, & Nevitt Sanford, 1950). In this study, it was observed that during the economic depression of the 1930s, the workers and have-nots in Germany did not revolt. They instead remained outside the political scene and were apathetic. Parts of the middle class, by contrast, actively engaged in politics, as did the cultural and economic elites. In light of these facts, the Frankfurt School criticized the Marxian notion concerning the direct effects of pauperization on political radicalism. The Frankfurt School introduced an original hypothesis by putting "personality" as an intermediate variable between social characteristics and political consciousness, between class and ideology. In their view, belonging to a deprived group does not lead to active political behavior but encourages the development of an authoritarian personality (i.e., subordination to authority, conformity to rigid social norms, rejection of those not sharing these norms). Instead, authoritarian personalities show deference to authorities and remain outside the political field.

More recent interpretations came from other studies completing the traditional authoritarian personality theories through a hypothesis proposed by Leo Festinger (1957) in *Theory of Cognitive Dissonance*: Awareness of the low likelihood of benefiting from an "achievement society" generates a mental state of anomie, social isolation, lack of confidence in the system, and apathetic behavior. Socioeconomic frustration and status anxiety thus mediate the correlation between class, authoritarian values, and anomie. From this tradition, it can be concluded that modern societies generate apathy in different ways but especially by creating gaps between aspirations, values, and realities. This body of literature has an obvious relationship with another intellectual tradition, which is known as the relative deprivation or relative frustration hypothesis. This hypothesis states that, despite the socioeconomic deprivation and social inequalities experienced by the popular classes during the periods of economic stagnation, there was no political protest. To explain this paradox, political sociology frequently makes reference to a Tocqueville effect. Alexis de Tocqueville proposed a theory of rising expectations that may offer some clues to the paradox of apathy and mobilization (those who normally would have interest to mobilize are not doing so). According to this theory, it is not deprivation but the obstruction of favorable prospects and rising expectations that encourage social protest and radical political behavior. As remarked by Van Snippenburg and Eisinga Scheepers (1991), citizens experiencing such frustrations "become morally indignant, sometimes socially envious, and they easily turn to political protest and radicalism as a consequence" (p. 45). This Tocqueville effect has been used to explain political protest when social groups experience significant discrepancies between their raising expectations, the realities of their occupational chances, and the possibilities for participating in political decision making.

More recently, the concept of apathy has been rejuvenated and updated by a set of notions proposed in the context of empirical studies on legitimacy and political support. Political apathy has been mentioned in the explanation of the apparent lack of political support and political trust in contemporary democracies. This also used to be the case with the concept of democratic deficit that was first proposed in the context of European Union (EU) integration and its legitimacy. Since the beginning of the 1990s, a significant body of literature has investigated the democratic gap in EU institutions. This democratic deficit is supposed to explain the lack of interest, motivation, and civic engagement of European citizens in the process of EU decision making. The idea of a democratic deficit or democratic disenchantment is also related to a set of phenomena identified as the emergence of critical citizens (Pippa Norris, 1999). According to Norris, critical citizens aspire to democracy and consider it as their ideal form of government, but at the same time they are deeply skeptical when evaluating how democracy works in their own country.

Different explanations have been proposed for such phenomena and are not fully integrated even if going in a common direction. Since the mid-1990s, for instance, many commentators pointed out that if contemporary democracies were not facing a system crisis, at least they were facing a syndrome of disenchantment and democratic malaise: Citizens in many industrial societies had become disengaged from public space. It is worth noting that the theory of critical citizens is perfectly compatible with the portrait of apathetic citizens or skeptical, even angry or cynical citizens rather than with the image of apathetic citizens expecting public services but not supporting the democratic system actively. The set of explanations coming from this literature about the democratic gap in contemporary societies claims that this phenomenon arises from some combination of growing public expectations, negative news, lack of public trust, and failing government performance. Complementary explanations have been proposed, such as social capital theory focusing on a decline of civic engagement and social trust (Robert Putnam, 1993) or postmaterialist theories (Ronald Inglehart, 1997), emphasizing the changing value orientations linked to processes of individualization. Indeed, the World Values Surveys document in a large set of countries that individualization processes detach citizens from the public sphere. But the postmaterialist revolution also shows the capacity of educated citizens to be the changing force of a cultural revolution in many societies, education playing a crucial role in the cognitive mobilization processes underlying such value changes.

A challenging view about apathetic citizens has been proposed by Pierre Rosanvallon, a French historian engaged in research on the intellectual roots of contemporary democracy. In the context of lamentations about the low turnout in European elections and the apathy of citizens regarding EU issues, Rosanvallon contradicted the myth of a passive citizen. His main argument was that the phenomenon of abstention reflects a mutation of, and not just a decline in, the public sphere. According to him, the decline in voter turnout has often been accompanied by other democratic activities. Rosanvallon substitutes the view on democratic fatigue in contemporary societies (a view that too many elections weaken the "democratic appetite" of citizen) with a distinction between three facets of democracy: (1) *expression* of citizens in making judgments on the rulers and their actions; (2) *involvement*, which encompasses all means by which citizens get together to produce a common world; and (3) *intervention*, which consists of all forms of collective action.

According to Rosanvallon, the characteristic elections should be superimposed on these different forms of civic life. In other words, voting is the most condensed form of democracy and the most organized and visible. Rosanvallon observes that although lower participation in elections has undoubtedly eroded the democratic expression, involvement and intervention have been strengthened. As an example, Rosanvallon reports the following data on the French case: In terms of involvement, participation has not declined. The French are now more likely to have signed petitions (68% in 2000 compared with 53% in 1990). They have also taken part in more events and demonstrations to make their voices heard. In other words, citizens' apathy does have some empirical foundations, but it is a multifaceted and multidimensional phenomenon. Depending on the definition, the available empirical indicators can change significantly the views about a supposedly apathetic citizenry.

Bruno Cautrès
CEVIPOF/Sciences Po
Paris, France

See also Alienation; Anomia; Dissatisfaction, Political; Legitimacy; Participation; Postmaterialism; Social Capital

Further Readings

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper & Row.

Almond, G. A., & Verba, S. (1963). *The civic culture*. Princeton, NJ: Princeton University Press.

Festinger, L. (1957). *A theory of cognitive dissonance*. Palo Alto, CA: Stanford University Press.

Inglehart, R. (1997). *Modernization and postmodernization*. Princeton, NJ: Princeton University Press.

Jahoda, M., Lazarsfeld, P. F., & Zeisel, H. (1971). *Marienthal: The sociography of an unemployed community*. Chicago: Aldine-Atherton.

Norris, P. (1999). *Critical citizens: Global support for democratic governance*. Oxford, UK: Oxford University Press.

Putnam, R. (1993). *Making democracy work*. Princeton, NJ: Princeton University Press.

Rosanvallon, P. (2006). *La contre-démocratie: La politique à l'âge de la défiance* [The cons-democracy: The policy at the age of mistrust]. Paris: Éditions du Seuil.

Van Snippenburg, L. B., & Scheepers, P. (1991). Social class and political behavior during a period of economic stagnation: Apathy and radicalism in the Netherlands, 1985. *Political Psychology*, *12*(1), 41–63.

Area Studies

Area studies is the generic term for multidisciplinary social research that focuses on specific geographic regions or culturally defined areas. The largest scholarly communities in this respect are loosely defined as Asian, African, Latin American or Middle Eastern studies, together with a variety of subfields (Southeast Asian studies, Caribbean studies, etc.). Political science plays a major role in area studies programs, which typically also draw on disciplines such as history, sociology, ethnology, geography, linguistics, literature, and cultural studies. In political science, the presence of area studies is strongest in comparative politics, but they are also incorporated in international relations (IR) and other fields. This entry first describes the origins of area studies, from early colonial studies to the development and institutionalization of area studies in the Cold

War era. It then outlines the not always easy relation between area studies and the discipline of political science, highlighting key achievements as well as remaining problems. The final section turns to the current research frontier, which is marked by a new impetus for comparative area studies and by transnational perspectives that question, among other things, established concepts of space or area.

The Origins of Area Studies

Today's area studies can be seen as having their origins in the colonial expansion of European powers and the accompanying academic efforts to better understand those "foreign" or "native" societies that demonstrated a variety of languages, cultures, and social organizations hitherto unknown. In this sense, area studies emerged as a "child of empire," often driven by commercial and political interests or the perceived "civilizing mission" of the colonial powers. At the same time, the study of ancient civilizations, ethnic codes, social hierarchies, or foreign languages was part of the much broader process of the extension of Western science across the globe. While from the mid-18th century, European capitals began to display the treasures and arts of "exotic" civilizations as much as of ancient civilizations in public museums, the 19th century saw the establishment of colonial studies in European universities. In the United States, interdisciplinary centers for area studies first emerged after World War I, and they received a strong impulse after World War II, parallel to the U.S. rise as a global power. A better understanding of societies in Asia, Africa, the Middle East, and Latin America was seen as urgent in the context of the Cold War rivalry between competing superpowers looking for local clients and supporters, particularly in the Third World. (A similar, security-driven incentive to promote the study of foreign cultures was again seen after the terrorist attacks of September 11, 2001.)

The work of German geographer Alexander von Humboldt (1769–1859) was a forerunner of area studies. Von Humboldt was also one of the founding fathers of multidisciplinarity, which then ranged from social to natural sciences. At a later stage, a critical strand of area studies emerged that openly condemned colonial practices. This branch emphasized respect for other cultures, challenged the supposed universality of the Western worldview and the "Eurocentrism" inherent in theories claiming general validity, and advocated mutual learning instead of unilaterally copying Western social or political models.

A common legacy of all strands of area studies, however, is that they almost always refer to "other" areas. There are no "German studies" in Germany or "U.S. studies" in the United States. Until today, area studies have focused predominantly on the non–Organisation for Economic Co-operation and Development (OECD) world. A remarkable trend, however, has been the emergence and expansion of "European studies" in the wake of the growing importance of the European Union.

Area Studies and Political Science: An Ongoing Debate

The relationship between area studies and the discipline of political science has not always been an easy one. Particularly in the United States, the eclectic use of changing paradigms by area specialists has been criticized. Single-case studies have been attacked by scholars interested in generalization. At times, area specialists have been expected to simply deliver the raw data for universal theories. Conversely, area specialists have accused generalists of superficiality, claiming that an in-depth knowledge of the particular language, culture, and social context is needed to produce adequate analyses of cases that do not conform to the OECD "standard model"—on the empirical realities of which most political science theory has been generated.

Contemporary political science has developed most strongly in the post–World War II period and in the transatlantic North. This explains why the bulk of existing theories and methods have a clear "transatlantic" or "OECD bias." Democracy, elections, state and administration, security, and so on are conceptualized in mainstream political science using idealized blueprints of Western societies. It cannot be denied that these concepts have been highly influential in other world regions as well and have been internalized by many scholars, elites, and citizens from Argentina to Zimbabwe. However, they are also frequently contested and compete with other organizing principles of society. The trajectories toward modernity have in fact varied enormously around the globe, with religion, family bonds, or industrialization playing divergent roles in different world regions. The exclusive reliance on "imported" models without local adaptation, so area studies scholars have argued, produces systematic errors in the comprehension of how "other" political systems function. In a more and more globalized world and for advocates of a practice-oriented approach in the subdiscipline of IR, it has become essential to understand the motivations and capabilities of actors in the most remote places of the globe in order to appropriately analyze security and economic challenges for their own societies.

Nonetheless, a balanced view of the strengths and weaknesses of the approaches of both mainstream political science and area studies has led to a more fertile relationship and compromises between "nomothetic" and "idiographic" worldviews. Some of the most prolific strands of research in political science have been strongly influenced by area specialists; for instance, the understandings of corporatism as well as of transitions to democracy in the 1980s were largely developed by Latin American studies scholars, before their concepts "traveled" to explain European and other cases. More recently, the topic of the consolidation and reform of "hybrid regimes" has been studied in Africa and South East or Central Asia, and much of the study of persistent authoritarian rule draws heavily on the scholarship of area specialists of the Middle East and Asia. Similarly, studies of clientelism or nationalism, informal politics or political culture, and development models or revolution would not be possible without the insights of area specialists. Some scholars go so far as to call "area studies" a misnomer, stating that area-focused scholarship by political scientists should be understood as an integral component of comparative politics.

The "Orientalism" Debate and the Demarcation of the "Areas" Studied

Criticism of area studies has also been raised from within the regions under scrutiny, most prominently in the "orientalism" debate kicked off by the 1978 publication of Edward Said's influential critique of Western constructions of the "Orient." Area studies, according to this critique, expressed an imperialist and condescending worldview regarding the "other"; thus, the object of research had to be redefined, and a complete overhaul of the production of academic research on non-Western societies was necessary. Postcolonial studies emerged from this line of thought as a competing paradigm of research that sharply criticized mainstream Western academic approaches as being part of an international system of domination in continuity with the colonial past. While strongest in literary theory and cultural studies, postcolonialist approaches also concern social and political science.

A different line of critique has come from scholars in countries of the global South, who consider themselves political scientists, just like any North American or European scholar studying his or her own country, but find their research on Brazilian parties, South African trade unions, or Indian class structure labeled as area studies rather than being accepted as political science on an equal footing with similar research in the OECD world.

A particular concern in area studies is the exact territorial demarcation of the "areas" under investigation—all the more so given the recent emphasis on transnational and transregional interrelationships. Is it appropriate that African studies more often than not deal exclusively with the Africa south of the Sahara? Put differently, is North Africa part of both African and Arab studies? What implications does the choice between "Arab world" and "Muslim world"—with an emphasis on ethnicity rather than religion—have for the understanding of the region? Does it make sense to group Southeast Asian, Central Asian, and South Asian studies together under the label of Asian studies? While Latin American studies encompass a region with a largely shared historical and cultural legacy, some recent initiatives have shifted to a more hemispheric approach under the label of American studies, which includes Canada and the United States.

Intellectual debates on these matters abound, but the persistence of the existing classifications is a sign that they continue to provide a basis for the production of meaning. Recent research has in fact argued that "regional context" is one crucial variable in many quantitative research designs, and in IR, the concepts of a "world of regions" and an emerging regional architecture in world politics have become an important focus of scholarly attention, implicitly strengthening the key assumptions underlying the concept of area studies.

Comparative Area Studies

A distinct and new approach is "comparative area studies." In addition to intraregional and interregional forms of comparison, the long-marginalized field of cross-regional comparisons has emerged as part of the agenda, for example the comparison of social security systems in Venezuela and India or of party systems in Tanzania and Indonesia. The case for comparative area studies is twofold: First, given that the emphasis on regional expertise has been their raison d'être, the different branches of area studies have developed largely separated from each other over decades. The ability of concepts "to travel" from one regional context to another has yet to be investigated in an appropriate way. While, for instance, the analysis of the paradox of concomitant democratization and the persistence of neo-patrimonial rule was pioneered in the African context, its general usefulness can be tested on Latin American and Asian cases. Such cross-area analysis allows the detection of area specificities as much as commonalities. A more general theory may emerge more easily from such a research design than from single-area approaches.

The second line of argument for comparative area studies holds that cross-regional small-*N* comparisons are particularly well suited to perform a mediating role between area-focused scholarship and the general theoretical and methodological debates in political science. While there will be few scholars with equally profound expertise in distant world regions, the demanding prerequisites for conducting such cross-area research can be systematically fulfilled through collaborative projects between scholars with different regional expertise but shared research interests. Area studies are not limited to a particular set of methods or schools but employ the wide array of methods used in the disciplines involved. Typically, however, case studies and paired and small-*N* comparisons have been dominant, and—given area scholars' emphasis on profound knowledge of the area under investigation—qualitative approaches have traditionally been stronger than quantitative studies.

Finally, the dynamic development and increasing self-assertion of scholarship from the regions that are the topic of area studies have led to growing sensitivity toward ethical considerations voiced in "the South." Area specialists in Northern countries are constantly pressed by their colleagues from Africa, Asia, Latin America, or the Middle East to share their privileged access to libraries, data, and funding. One element of a more collaborative approach is, for example, the open-access movement to publish social science findings free of charge on the Internet for readers worldwide. This may eventually narrow the gap in the access to relevant current literature between the "Harvards" and the "Have-nots."

Andreas Mehler and Bert Hoffmann
German Institute of Global and Area Studies
Hamburg, Germany

See also Comparative Methods; Hybrid Regimes; Neo-Patrimonialism; Orientalism; Postcolonialism

Further Readings

Acharya, A. (2006). *International relations and area studies: Towards a new synthesis?* (State of Security and International Studies Papers No. 2). Nanyang, Singapore: Institute of Defence and Strategic Studies.

Basedau, M., & Köllner, P. (2007). Area studies, comparative area studies, and the study of politics: Context, substance, and methodological challenges. *Zeitschrift für Vergleichende Politikwissenschaft, 1*(1), 105–124.

Harbeson, J. W., McClintock, C., & Dubin, R. (Eds.). (2001). "Area studies" and the discipline: Towards new interconnections. *PS: Political Science & Politics, 34*(4), 801–804.

Said, E. (1978). *Orientalism. Western representations of the Orient*. New York: Pantheon.

Sil, R. (2009, Fall). Area studies, comparative politics, and the role of cross-regional small-*N* comparison. *Qualitative and Multi-Method Research*, pp. 26–32.

Szanton, D. L. (Ed.). (2004). *The politics of knowledge: Area studies and the disciplines*. Berkeley: University of California Press.

Arms Race

An arms race may be defined as a pattern of competitive acquisition of military capability between two or more countries. The term *arms race* is often used quite loosely to refer to any military buildup or spending increases by a group of countries. This definition requires that there be a *competitive* nature to this buildup, often reflecting an adversarial relationship. The arms race concept is also used fruitfully in other fields, especially evolutionary biology; however, this entry deals only with military arms races.

Examples of Arms Races

One example of an arms race is the "Dreadnought" arms race between Germany and Britain prior to World War I. In the early 20th century, Germany as a rising power sought to challenge the United Kingdom's traditional naval dominance. In 1906, Britain launched a new, more advanced warship, HMS *Dreadnought*, triggering a naval arms race. Between 1909 and the outbreak of World War I in 1914, Britain launched a further 19 "dreadnoughts" and a further nine battle cruisers, while Germany launched 13 dreadnoughts and five battle cruisers. This arms race is often cited as one of the causes of World War I.

The Cold War nuclear arms race between the United States and the Soviet Union is another example of a 20th-century arms race. The United States' use of nuclear weapons to end World War II led to a determined effort by the Soviet Union to acquire these weapons, leading to a long-running nuclear arms race between the two superpowers. The Soviet Union conducted its first nuclear test in 1949. At the end of 1956, the United States had 2,123 strategic warheads to the Soviet Union's 84. These numbers increased rapidly over the subsequent 30 years. The U.S. arsenal peaked in 1987 at 13,002 warheads, the Soviet Union 2 years later at 11,320 (figures from Stockholm International Peace Research Institute). The end of the Cold War effectively ended this arms race.

Arms races may involve a more general competitive acquisition of military capability. This is often measured by military expenditure, although the link between military expenditure and capability is often quite weak. Such more general arms races are often observed among countries engaged in enduring rivalries, which may sometimes appear to follow each other's military spending levels, especially during periods of heightened tension. Examples of such arms races include India–Pakistan, Israel–Arab states, Greece–Turkey, and, since 1991, Armenia–Azerbaijan.

Consequences of Arms Races

Arms races are frequently regarded as negative occurrences in both economic and security terms.

Large-scale arms acquisitions require considerable economic resources. If two countries spend large sums of money just to cancel out each other's efforts, the expenditure might well be seen as wasted. There is, however, considerable debate surrounding the economic effect of military spending. Some argue that it provides benefits through technological spin-offs, job creation, and infrastructure development; others argue that it displaces more productive forms of investment, while its final output is not itself productive. Certainly, countries that must import arms will see more negative economic effects of an arms race, and arms imports are a major contributor to Third World debt. Even for arms-producing countries, excessive military expenditure is likely eventually to have negative economic consequences. The Soviet Union's economic difficulties were certainly exacerbated by the very high proportion of the gross domestic product devoted to the arms race.

The question of whether arms races contribute to the outbreak of war is also the subject of considerable debate. An arms race may heighten fear and hostility on the part of the countries involved, but whether this contributes to war is hard to gauge. Some empirical studies do find that arms races are associated with an increased likelihood of war; however, it is not possible to say whether the arms race was itself a cause of war or merely a symptom of existing tensions.

One may also consider the gains for a country that "wins" an arms race in the sense of gaining a decisive military advantage. Arguably, the collapse of the Soviet Union, which left the United States as the sole global superpower, was partly due to the cost of attempting to keep up with the United States.

Modeling Arms Races

There is an extensive literature modeling arms races theoretically and empirically. These include game-theoretic models based on the "prisoners' dilemma" (PD), dynamic mathematical models based on the Richardson model, and economic models frequently based on a "utility-maximizing" framework. There is overlap between these categories.

Prisoners' Dilemma Models

The famous PD game is frequently applied to arms races between two countries. It is assumed that each country has a choice between a high or low level of arms. Each country's most preferred outcome is assumed to be where they choose high arms and their rivals low, gaining a clear military advantage. Their least preferred outcome is the reverse. However, the second best outcome for each is where both choose low; if no advantage is gained, it is assumed to be cheaper and more secure to avoid the arms race.

Each player's "dominant strategy" is to choose high arms, as whichever choice the rival has made, they do better by choosing high than low. This outcome illustrates the Nash equilibrium (named after John Nash), in which each player has chosen the optimal strategy in a noncooperative situation or zero-sum game, given the other player's strategy. The Nash equilibrium outcome of the game is therefore that both choose high. It is an equilibrium because neither player would change their own choice, given the choice of the rival. However, this arms race outcome is worse for both players than if both had chosen low arms. The logic of their rivalry traps them in a mutually disadvantageous situation.

However, the picture may not necessarily be so bleak, as in reality the "game" is not played once and for all but is an ongoing series of decisions, which can be modeled by the "iterated prisoners' dilemma" (IPD), where the PD game is played repeatedly by the same players. This opens the possibility for cooperation to emerge through reward and punishment strategies such as "tit for tat"—start by choosing low arms but then match the strategy chosen by the other player in the previous round. Experiments have shown this to be a highly successful strategy in IPD games.

While greatly simplifying real-world situations, the PD may be a useful metaphor to capture the essential dilemma facing countries engaged in an enduring rivalry.

The Richardson Model and Its Elaborations

In his seminal work *Arms and Insecurity*, Lewis Fry Richardson proposed a model (which he applied to the *Dreadnought* race) of an arms race between two countries where each country sets its military expenditure (or arms acquisition) level in each period based on its own and its rival's level in the previous period in an "action–reaction" pattern. This is modeled by the following equations:

$$M1_t = -a1M1_{t-1} + b1\mathrm{M}2_{t-1} + g1,$$

$$M2_t = -a2M2_{t-1} + b2M1_{t-1} + g2.$$

Here, $M1_t$ and $M2_t$ refer to the military spending levels of Countries 1 and 2 in years t and $t - 1$, respectively. The coefficients $a1$ and $a2$ (assumed to be positive) are "fatigue" coefficients, representing the difficulty of maintaining high levels of military spending. The coefficients $b1$ and $b2$, also positive, are "reaction" coefficients, measuring the tendency for each country to respond to the military spending of their rival, while $g1$ and $g2$ are autonomous "grievance" or "ambition" terms, representing each country's desire for military capability apart from the rivalry. Depending on the relative size of the fatigue and reaction coefficients, the arms race can either reach a stable equilibrium or spiral out of control.

The basic Richardson model has been extensively developed by other authors, both theoretically and empirically. Developments include taking into account the stock of weapons of each country as well as the rate of spending, introducing explicit economic criteria, and modeling the strategic dynamics of the relationship. Richardson models can easily be applied to empirical military spending data, using regression analysis to estimate the parameters of the equation for a pair of countries. The key question is whether the reaction terms $b1$ and $b2$ are significantly greater than zero—if so, an "action–reaction" or Richardsonian arms race is said to exist. A wide variety of theoretical models and statistical techniques starting from the Richardson framework have been applied to various pairs of countries.

Economic Models

A third approach is to assume that countries' military expenditure decisions are the outcome of an economic resource allocation process whereby the government seeks to achieve a set of economic, political, and security objectives by allocating spending between military and civil sectors. Neoclassical rational choice models are most frequently employed, but others are also used. When two rival countries are considered, each country's level of security is assumed to depend on both its own level and its rivals' level of military spending. Each country makes its decision taking into account the likely response of the other. The resulting models are similar to certain elaborations of the Richardson model: While one starts from the arms race and builds in economic allocation issues, the other starts with the allocation problem and builds in the rivalry.

Weaknesses of Richardsonian Models

While the Richardson model has proved extraordinarily fruitful in generating scholarly analysis of arms races, the Richardson approach has a rather poor empirical record in demonstrating the existence of actual arms races. Some, though not all, studies of India and Pakistan have found evidence of a Richardsonian arms race, but few other enduring rivalries have produced such empirical results and none consistently.

A problem of the Richardson model is that it assumes that the parameters of the relationship (the values of a, b, and g) remain constant, whereas in reality they may change over time depending on the changing relationship between the countries. A second is that it is most applicable to pairs of countries where the rivalry is the overwhelming factor for each country's security, which is relatively rare.

One suggestion is that the changing levels of tension or hostility between countries may be a better way of explaining their military spending decisions than the Richardson action–reaction framework—or perhaps a combination of levels of hostility with the rival's military spending. This would suggest that arms races are characterized more by short bursts of rapidly increasing spending during periods of high tension than by long-term, stable relationships between their levels of spending.

Sam Perlo-Freeman
Stockholm International Peace Research Institute
Solna, Sweden

See also Arms Race; Budgeting, Rational Models; Cold War; Game Theory; Models, Computational/Agent-Based; Prisoners' Dilemma; Rational Choice; Regression; Statistical Significance; Superpower; War and Peace

Further Readings

Anderton, C. H. (1989). Arms race modeling: Problems and prospects. *Journal of Conflict Resolution, 33*, 346–367.

Axelrod, R. (1984). *The evolution of co-operation*. New York: Basic Books.

Dunne, J. P., Perlo-Freeman, S., & Smith, R. (2008). The demand for military expenditure in developing countries: Hostility versus capability. *Defence and Peace Economics*, *19*, 293–302.

Dunne, J. P., & Smith, R. (2007). The econometrics of military arms races. In K. Hartley & T. Sandler (Eds.), *Handbook of defence economics* (Vol. 2, pp. 913–940). Amsterdam: North Holland.

Norris, R., Fieldhouse, R., Cochrane, T., & Arkin, W. (1991). Nuclear weapons. In *SIPRI Yearbook 1991: World armaments and disarmament* (pp. 3–40). Oxford, UK: Oxford University Press.

Ram, R. (1995). Defense expenditure and economic growth. In K. Hartley & T. Sandler (Eds.), *Handbook of defence economics* (pp. 251–273). Amsterdam: North Holland.

Richardson, L. F. (1960). *Arms and insecurity: A mathematical study of the causes and origins of war*. Pittsburgh, PA: Boxwood Press.

Von Neumman, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Attitude Consistency

Three types of attitude consistency can be distinguished: (1) consistency among the various components of an attitude, (2) consistency between different attitudes, and (3) consistency between attitudes and behavior. This entry addresses the relevant conceptualizations, measurements, and explanations.

Intra-Attitudinal Consistency

An attitude is said to be consistent when all elements elicit similar evaluative judgments. Intra-attitudinal inconsistency arises when some of these elements are positive in nature, while others are negative. Within the cognitive component, accessible beliefs may conflict, cognitions and affects may also conflict, and the various affects may differ, simultaneously. Intra-attitude consistency can be examined by comparing how respondents feel, how they think, and what their summary evaluation is. The many studies that showed low correlations and consistency between the cognitive, affective, and conative (behavioral) components have raised doubts about the early theorists' multicomponent view of attitudes. Martin Fishbein and Icek Ajzen restricted the term *attitude* to a person's evaluation and drew a clear distinction between cognitions, attitudes, behavioral intentions, and behaviors. From their point of view, an attitude is a person's general feeling of favorableness or unfavorableness and an index of the degree to which a person likes or dislikes an object. This one-component view of attitude opened the way to the development of more sophisticated explanatory models in which the summary evaluation is at the core of an attitude, while behavior or the intention thereto is an effect, and cognition (knowledge and beliefs) and affect (general moods and specific emotions) are origins of the attitude. The affects underlying attitudes are more easily accessible in memory than the cognitions, and they tend to predominate among ambivalent respondents with different evaluative-cognitive and evaluative-affective consistencies. Individuals differ in their tendency to base their attitudes on cognition or affect, as Geoffrey Haddock and Mark Zanna have shown. Consistent attitudes were found to be stronger, more predictive of subsequent intentions and behavior, more resistant to persuasive communication and change, and thus more stable than inconsistent attitudes. Inconsistent attitudes seem to result in more information processing and knowledge acquisition.

Interattitudinal Consistency

Attitudes toward the same object and attitudes toward different objects are said to be consistent when they do not contradict each other. Interattitudinal consistency with respect to the same object (dual-attitude model) is empirically examined by the extent to which a respondent's explicit and implicit or habitual attitudes are related to one another. Interattitudinal consistency with respect to different objects is empirically examined by the extent to which these attitudes are related to one another and/or the level of correspondence between the respondent's attitudes and his or her ideological self-identification (e.g., consistent liberal or conservative attitudes), left–right self-placement, and respondent's value system and/or the number of issues on which a respondent adopts his or her (liberal or conservative) party's position. Interattitudinal consistency is considered desirable from the point of view of "correct" voting in representative democracies, requiring citizens to make well-founded voting decisions. Scholars vary in their views on the level of interattitudinal consistency. Philip Converse was one of the first who showed that the political attitudes that many people have are not consistent and referred to this aspect of public opinion as "nonattitudes." However, the evidence presented by Max Kaase and Kenneth Newton, using Eurobarometer data, suggests that public opinion (in Western Europe) is relatively internally consistent (at both the aggregate and the individual levels). Mass opinion about the welfare state and the scope of government is structured into relatively few attitude publics. Mass attitudes about spending on government services and attitudes toward policy issues cluster into three groupings of related issues (welfare, security, and quality of life). There is also a match between saying that a particular issue has high priority as a public service and expressing a willingness to pay taxes for it. Finally, the statistical associations with left–right orientation and party identification are usually stronger than with any of the social, economic, and demographic variables. However, this left–right factor generally explains only a small proportion of the variance in issue positions. A factor that can also help us understand the structure of political attitudes is values; there is now a substantial amount of evidence that values are a major source of structure for political attitudes, according to Stanley Feldman.

Explanations of Attitude Consistency

Consistency and inconsistency have both interpersonal and intrapersonal sources. Political socialization theory, represented by, among others, David

Easton, Kent Jennings, Richard Niemi, Judith Torney, and Russell Farnen, emphasizes interpersonal and contextual factors and the level of consistency or inconsistency of the cognitive and affective messages that people receive from the various socialization agencies (including political leaders, mass media, friends and colleagues, religious institutions, school, family) in particular. The scale runs from very high message consistency in totalitarian regimes and autocracies, via moderate consistency in representative democracies with weak elite polarization, to very low consistency in democracies with strong elite polarization.

Various intrapersonal factors have been identified. Individuals differ in their chronic tendency to engage in evaluative responding (which can be measured by a set of items developed by Blair Jarvis and Richard Petty). People also differ in their attempts to reduce or avoid the tension that they experience when they become aware of contradictions. To avoid the anticipated tension, people prefer exposure to consonant media and messages (e.g., Fritz Heider's balance theory and Leon Festinger's cognitive dissonance theory). The level of political knowledge is another important variable. In general, the mean level is very low but the variance is very high. The higher the level of political knowledge, the more tightly structured or "constrained" the individual's system of attitudes, as Converse has shown. A few theoretical frameworks have been developed as attempts to explain the organization of political and social attitudes. Converse's "ideological constraint theory" holds that individual attitudes become consistent thanks to deduction from higher order orientations such as ideology. People differ in attitude consistency because they differ in the use of ideological schemas. Howard Lavine and Cynthia Thomsen have developed the "shared-consequences model" to understand the psychological mechanisms through which structural organization and consistency between attitudes develop. The core of the model is the strength of the perceived implicational relationship between a pair of attitude objects. Perceiving two policies, for example, legalized abortion and capital punishment, as instrumentally influencing a similar set of consequences, for example, the attainment of similar sets of value-related goals, constitutes a strong implicational relationship between them, which, in turn, creates consistency pressures. Consistency, then, is the psychological agreement between the valences of a pair of attitudes and the valence of the implicational relationship perceived to exist between them. The development of interattitudinal organization may proceed through a series of distinct stages: Interattitudinal connections may initially be derived from affective associations that issues have in common, next to self- and group-interest considerations, then from values, and finally from ideology.

Attitude–Behavior Consistency

Attitudes are studied primarily because of their expected effects on behavior. It is expected that the more favorable an individual's attitude toward some object, the more likely the individual will be to perform any given positive behavior and the less likely he or she will be to perform any negative behavior with respect to the object. In Allan Wicker's analysis of 45 studies, however, only a small minority of studies showed a strong or modest positive relationship between attitude and behavior. Hans Benninghaus's analysis of 57 studies supported Wicker's findings. Inconsistency between attitude and behavior is higher if the attitude object is not related to one's self-interest, if intra-attitude inconsistency is high, and if one's attitude conflicts with the subjective norm or the perception of how relevant others think one should behave.

In an attempt to improve the attitude–behavior link, the conceptualizations and operationalizations of the two key concepts have been reconsidered. Fishbein and Ajzen introduced the conceptual restriction of attitudes to a person's evaluation or feeling of favorableness or unfavorableness. With respect to behavior, distinctions were introduced, for instance, between behavioral categories and single actions, isolated individual behavior and behavior in the presence of others, and behaviors under much or little volitional control. In subsequent behavior-explanatory theories, attitude is still an independent variable, though not the one and only, the most important, or the most immediate variable. In Fishbein and Ajzen's "theory of reasoned action," which presents a causal chain to predict behavior, the immediate determinant of behavior is the individual's intention to perform the behavior in question. Attitudes toward the behavior and the "subjective norm" are the two determinants of the individual's behavioral intention. Thus, behavioral and normative beliefs are the antecedents of attitude and subjective norm and, as a result, ultimately determine intention and behavior. This reflects a cognitive approach to explaining attitudes and a rational view on human behavior. One of the main comments on the reasoned-action theory was that this theory is only applicable to purely volitional behavior. Ajzen's "theory of planned behavior" extends the theory of reasoned action by adding the perception of control over behavior. Perceived behavioral control is also cognitive and rational since it is the individual's set of beliefs about the presence or absence of requisite resources and opportunities. New meta-analyses by, among others, Min-Sun Kim and John Hunter and Christopher Armitage and Mark Conner, showed stronger attitude–behavior relationships and a relatively strong predictive power of the theory of planned behavior. Many studies have demonstrated the applicability of the theory in various content domains, including the political. In Marco Perugini and Mark Conner's "model of goal-directed behavior," the basic variables of the planned behavior theory were added, together with three other key variables: desire/motivation, emotions, and habit.

In general, affective variables are gaining in importance in theory and research aiming at explaining the consistency between attitudes and behavior. Promising political science theories, building on new insights from political neuroscience, are George Marcus's "affective intelligence theory," Milton Lodge and Charles Taber's "theory of motivated political reasoning," and Richard Lau and David Redlawsk's "model of motivated reasoning."

Henk Dekker
Leiden University
Leiden, Netherlands

Jolanda van der Noll
Bremen University
Bremen, Germany

See also Attitudes, Political; Beliefs; Conservatism; Liberalism; Meta-Analysis; Political Psychology; Political Socialization; Values

Further Readings

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*(1), 27–58.

Dalton, R. J., & Klingemann, H.-D. (2007). *Oxford handbook of political behaviour*. Oxford, UK: Oxford University Press.

Dekker, H., Dijkgraaf, B., & Meijerink, F. (2007). Behaviour and attitude: Effect of attitude on behavioural desire with respect to foreign countries and peoples and Germany and Germans in particular. In E. Poppe & M. Verkuyten (Eds.), *Culture and conflict: Liber Amicorum for Louk Hagendoorn* (pp. 197–215). Amsterdam: Aksant.

Neuman, W. R., Marcus, G. E., Crigler, A. N., & MacKuen, M. (Eds.). (2007). *The affect effect: Dynamics of emotion in political thinking and behavior*. Chicago: University of Chicago Press.

Sears, D. O., Huddy, L., & Jervis, R. (Eds.). (2003). *Oxford handbook of political psychology*. Oxford, UK: Oxford University Press.

ATTITUDES, POLITICAL

Individuals hold attitudes—and eventually change them—in relation to a wide diversity of topics. They eventually establish connections between one topic and another. Thus, for example, for some individuals, attitudes toward a political topic may be based on attitudes toward other subjects not related to politics. The idea that political attitudes constitute, or should constitute, a separate field has two main roots. One is the tradition connecting the sphere of politics with the kind of well-articulated sets of ideas frequently called "ideologies"; the other is the normative tradition postulating that human beings should approach the realm of politics in a way that is different from how they normally approach other spheres of life, basically by putting aside impulses, affects, and passions. We hear and record opinions, but no one can hear or see attitudes. Attitudes lie within individuals' minds; therefore, they are nonobservable, or latent, properties. To speak about "attitudes," we need to build, analytically, the aspects of reality the existence of which we postulate. The concept of attitude as it is employed in various political contexts and the factors shaping them are discussed as follows.

Definition and Measurement

A critical problem, as ancient as the notion of "attitudes" itself, is how to measure them. Being, by definition, unobservable aspects of reality, any attempt to measure them is based on assumptions. The first and best known attempts to measure attitudes employed projective tools, scales, attitudinal indices, and mathematical models allowing inferences about latent structures through manifest observations and data. All of these tools are based on the principle of a stimulus (a question) and a response (a verbal answer) that elicits some unobservable mental state. The idea that attitudes can be measured by recording manifest verbal expressions from which they can be inferred had a significant impact on the development of the techniques of quantitative survey analysis. This concept provided a logic, a methodological and an empirical support for the field of opinion research.

Let us think of voting. How do we account for the particular decision each voter makes at the ballot box? Is this act something determined from outside himself or herself, such as his or her class position, or is it derived from some kind of "objective interest"? Is it a random impulse? Or is it a consequence of the multiple messages an individual is exposed to in a communication process called an electoral campaign? Or is it to a great extent the result of media and public relations manipulation? Or is it the result of influences and pressures generated in the social groups the person belongs to—family, the neighborhood, friends, colleagues, and peer groups? Is the vote the product of political convictions and ideas of an individual and his or her ability to match those ideas with the ones the candidates communicate? Or is it just the result of preferences on any of the attributes, or issues, the voter sees in a candidate?

However, political convictions or preferences cannot be directly observed. They can only be inferred from what a particular voter says. This nonobservable, latent, level is called *attitude*: the propensity to produce certain behavioral responses relating to specific subjects, once the individual has been exposed to a stimulus. To the extent that such attitudes determine opinions and behavior, the individual is an autonomous actor rather than a passive object moved by causes external to him or her.

Attitudes or *predispositions*, as they are frequently called, vary from one individual to another. For some individuals, "ideas," in the sense of a set of articulated beliefs, matter; to others, tastes are more important. For some individuals, particular issues are more relevant than general concepts; for others, moral values are more important than particular preferences. So, some people may vote because of the positions of candidates or parties on specific issues, whereas others may vote with regard to more general public policies. Some individuals prefer a candidate just because they like or trust him or her, or they base their choices on the perceived honesty of the candidate irrespective of his or her standing on public policy issues.

The idea that there are voters who make their decisions in an erratic, random way has not been alien to the social sciences. Frequently, this idea implies the assumption that individuals are strongly dependent on mass media, leaders, or material incentives, so they are not really autonomous actors. In any case, *attitudes* are important factors accounting for individual behavior and thought. They constitute a term in an equation intended to describe why autonomous persons do the things they do, other terms being *information* coming to the individual from society and *pressures and influences* from the near environment.

The standard representation of attitudes postulates the existence of levels of generality and stability in the judgments people are able to formulate about themselves and the external world. A higher degree of generality means that an idea refers not to a single or particular subject but to a wider set of topics; a higher degree of stability means that ideas sustained through time resist the impact of new information the individual is continually exposed to. In these assumed hierarchies of judgments each individual holds in his or her mind, there are, on one extreme of the continuum, the more general and stable ones—values, values, ideologies; on the other pole, judgments refer to particular issues—opinions; and in the middle are judgments that are less general and stable than the first ones but not as occasional or particular as the latter—attitudes. The expression *belief system* is sometimes used to refer to the relationship between the different attitudes a person can hold in a stable manner.

The term *attitude* denotes an individual's propensity to think and to form opinions in a stable way when he or she refers to a specific topic. "Attitude" also refers to a propensity in the individual to mobilize certain mental contents, stable points of view, every time he or she formulates specific judgments, utters opinions, or decides on certain conducts.

Citizens do have predispositions toward public affairs, governments, and elections. These vary among individuals. For example, some individuals hold conservative points of view, others hold liberal ones; some are nationalist, others universalist; some favor a certain public policy, others do not.

Some dimensions of predispositions are thought of as if they were a given, by a sort of necessity, in each individual. Thus, it is frequently assumed that individuals are always either liberal or conservative, either left wing or right wing. Empirical evidence in many countries does not support such assumptions. Sometimes, the evidence instead suggests that there are individuals who to some extent may be both liberal and conservative, depending on the issues that are considered. In some countries, people define themselves as oriented to the left, right, or center; in other countries, these concepts are not relevant for many people.

Attitude Formation

How these predispositions are formed is a matter of debate among social scientists. Predispositions may result from genetic factors, socialization, learning, or many other influences. Two main traditions have developed within the social sciences. One places emphasis on *interests*: Alexis de Tocqueville said that people's opinions change with their fortunes. The other focuses on *membership groups* each individual takes part in: Émile Durkheim talked of "likenesses of consciences" in contrast to the effects of the division of labor. The idea that each social position must correspond to a particular set of attitudes is deeply rooted in social theory; it frequently receives empirical support, but the explanation to account for that correlation is far from being consensual.

The theory of attitudes postulates an inferred, nonobservable relationship between a given attitudinal dimension, which is by definition unobservable, and a manifest, observable one. We can conceive of that relationship as a probabilistic one, in the sense that there is no perfect link between the real position of a person on the attitudinal dimension and his or her manifest answer to a particular question at a given time. An individual can be at a certain latent state in the dimension of "nationalism," but that position does not determine a single answer the individual would produce to a given question aimed to elicit that attitude. A student who knows the content of a given topic very well is supposed to answer all the questions in an exam correctly, but sometimes he or she fails to do so.

We speak, for example, about attitudes toward the government. We could conceive of these attitudes as ranging from a very favorable attitude to a very unfavorable one. To measure that attitude in a particular individual at a given time, we may ask him or her: "What do you think about the performance of the present government?" and we possibly offer him a limited set of options: "very well," "fairly well," "about average," "fairly bad," or "very bad" or 5, 4, 3, 2, 1, where 5 is a *very favorable* and 1 a *very unfavorable* evaluation. The relation is probabilistic, because, even if the individual usually is in favor of the current government, some days he or she will not necessarily give the same answer, assuming we ask the same question in different ways and that he or she does not remember the previous answers. Moreover, different researchers can construct the scale in different ways: Some offer five options, as in the example, while others offer four or seven choices. The answers generate a variable that is statistically dependent on the latent one, but that variable could also be dependent on many other variables, most of them probably unknown to us.

Critics of this method of measuring attitudes may argue that all the other factors that may influence a response, including the more occasional or situational ones, could reduce the relationship between the latent dimension and the manifest one to zero, making the measurement irrelevant and not valid to measure what it intends to measure. The supporters of the method think that the range of variation between these two levels of reality is far more restricted and that there is no better way to approach attitudes—at least in a quantifiable way.

It is commonly accepted that political attitudes help the analyst explain and predict political preferences and behavior. There is a great deal of research examining the effects of electoral campaigns, political communication, and messages delivered by political leaders on voters' attitudes. While there are no universal generalizations supported by empirical research, some firm trends can be established. For example, a voter's level of education and the quantity of information an individual may acquire facilitate a higher propensity to change; membership groups also have an effect, but this effect tends to become weaker the more complex societies are. Political parties in democratic systems used to have a strong influence on voters, but this influence is becoming weaker as the number of citizens who feel independent from parties is increasing. The role of ideology in the propensity of individuals to become exposed to certain kinds of messages and to be receptive to them is also a matter of discussion. The influence of the print media is also a matter of widely divergent points of view—that is, whether the media influence public attitudes or whether they adapt themselves to the prevailing attitudes of their publics.

We know that many citizens today hold stable political predispositions and do not change them easily. They derive their opinions and particular political decisions from those predispositions. However, they do not link their predispositions to particular opinions and decisions in a strongly linear way. Rather, they afford a certain degree of variability in the opinions they form at every moment—including the voting decision. So, different persons eventually may make the same voting decision for different reasons, while persons with similar attitudes may eventually arrive at different voting decisions. Electoral campaigns are sources of intense stimuli, concentrated within a short period of time, which voters cannot easily resist. During a campaign, political advertising, press information, and the messages of the candidate and other leaders are able to produce some variability in particular voting behaviors even while voters' attitudes may remain relatively stable. Sometimes, those intense messages are able to modify attitudes as well.

Some citizens decide their vote by taking into account both their own attitudes and the ones they attribute to the candidates and then matching them. But there are others who tend to take into account "tactical" considerations, such as defining their vote against a candidate rather than in support of another or to voting for the presumptive loser in order to prevent the winner from obtaining a landslide. To make that kind of decision, electoral polls may influence some voters.

The principle that political behavior, and specifically voting behavior, must be related to what citizens have in their minds is at the root of democratic theory. Quite often, mass democracy is judged negatively precisely on the assumption that voters lack the capacity to make consistent judgments anchored in adequate information. Thus, the ancient problem of the relationship between attitudes and behavior is central in the realm of politics—be it formulated in these terms or not. Even if attitudes, in whatever way they have been formed, tend to be stable and do not vary at random, they are susceptible to change. One important question is whether human beings are able to modify their attitudes in an autonomous way—that is, by thinking and reasoning themselves—or whether they change their attitudes under the influence of messages from the outside world.

Attitude Change

We are far from a theoretical consensus on how attitudes can change. While one theoretical perspective postulates that better informed, more consistent persons are more stable in their attitudes, another theoretical perspective proposes that they are more susceptible to nuances and variations in particular opinions. Recent research tends to give support to the view that particular opinions may vary, to some extent, independently from well-established attitudes. Individuals have many different ideas in their minds; so, they may hold different attitudes, some of them more firmly rooted than others. The assumption that more general judgments are more stable in the mind of individuals does not imply that specific opinions and decisions may vary to a greater extent. That does not necessarily mean that citizens are inconsistent but, rather, that they are flexible.

When an individual receives information from outside, he can dismiss it or accept it; in the latter case, he can process it and record it, totally or partially, in his memory. The new information may have different effects on already existing attitudes: It can modify some of them, it can be conserved without modifying previous attitudes (sometimes even if it produces some inconsistencies), or it can be rejected. The effect of a new piece of information on previous attitudes can also be sometimes delayed: The new information has been taken into account, but the mind works on it slowly. So there can be a lag between the moment a stimulus is produced and the moment an effective change of attitude takes place. It is possible that those effects are more intense on some attitudes held by an individual than on others; therefore, as each individual holds a collection of interrelated attitudes rather than many separate ones, it is possible that some attitudes change while others do not, thus producing different patterns of consistency or inconsistency over time.

Let us take the example of electoral campaigns. Some voters do not change their attitudes under the effect of a campaign while others do; but to some voters, the effect comes with a lag. So, the transition from one attitudinal property to another could be nil, immediate, or deferred—that is, some persons may adopt a new behavior or a new opinion after a period of time. This leads to the notion of a *probability of transition* from one attitudinal state to another as a function of a new stimulus or new information received. Much of public opinion research focuses on these effects, frequently analyzing them by means of statistical association coefficients and contingency tables.

In summary, attitudes may change as the result of changes in the social position of persons or of changes in the sociocultural context or because of the effect of communication—press media, advertising, electoral campaigns, leaders' messages.

To account for variations in voting behavior, it is useful to transpose the electoral market to any situation where individuals must make a choice between different options. In those situations, there are factors that condition the individual's own perspective and there are factors that are under the control of external sources. When the individual who has to make a choice is exposed to political communication, he or she may or may not be motivated by well-reasoned political purposes. There is no rule about why and how individuals process their own motivations and the information they receive from outside.

A model of the effects that lead to voting decisions in each individual citizen may take into account three forces: (1) the initial position of each individual, (2) the various forces that push him or her toward a change in one direction, and (3) the various forces that push him or her in the opposite direction. The balance of those three types of forces determines whether the person will finally vote for the candidate he or she would vote for if the campaign did not exist and other factors remain the same or whether he or she changes his or her voting decision. Political campaigns are dynamic factors moving against the other factors that keep the system stable.

Generally speaking, attitudes are at the base of the demand side; they help citizens form expectations and respond to what they are offered. Traditionally, political studies have given more emphasis to the supply side: articulated ideologies and widely shared systems of ideas, political parties, leadership, mass media. More recently, the focus has gradually shifted to the demand side; there has been increased interest in understanding individual predispositions and attitudes and the individual sources of opinions and behavior. Market research, motivational research, opinion research, and communications research have been the engines of the approaches centered on the demand side.

Manuel Mora y Araujo
Universidad Torcuato Di Tella
Buenos Aires, Argentina

See also Attitude Consistency; Beliefs; Electoral Behavior; Electoral Campaigns; Public Opinion; Survey Research

Further Readings

Converse, P. (1964). The nature of beliefs systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 206–261). New York: Free Press.

Converse, P. (1975). Public opinion and voting behavior. In F. Greenstein & N. Polsby (Eds.), *Handbook of political science* (pp. 75–169). Reading, PA: Addison-Wesley.

Mora y Araujo, M. (2007). *The power of conversation.* Buenos Aires, Argentina: La Crujía/Ipsos.

Noelle-Neumann, E. (1984). *The spiral of silence.* Chicago: University of Chicago Press.

Price, V. (1992). *Communication concepts: Vol. 4. Public opinion.* Newbury Park, CA: Sage.

Sniderman, P. M., Brody, R. A., & Tetlock, P. E. (1991). *Reasoning and choice: Explorations in political psychology.* Cambridge, UK: Cambridge University Press.

Zaller, J. R. (1992). *The nature and origins of mass opinion.* Cambridge, UK: Cambridge University Press.

Audit Society

The concept of "audit society" refers primarily to a logic at work in regulatory and managerial processes and to tendencies, rather than to an epochal state or condition. In this context, an audit refers to a systematic assessment of institutional performance against agreed criteria and standards. The processes used to conduct such audits reflect the transformation of audit and evaluation practices into normative principles of organizational control with a wide reach. This transformation is in turn a function of what political scientists have called the "regulatory state," meaning in part a state with an operating philosophy of acting "at a distance" on policy domains, via instruments such as accounting and audit. Below, the role of the "audit society" idea as a critical resource is developed, followed by a number of methodological issues and concerns related to its potential application.

The idea of "audit society" provided an organizing motif for criticisms of the public service reforms in the United Kingdom (UK) and elsewhere during the 1980s and 1990s, known collectively as the "new" public management. While the original formulation of the idea relied largely on a synthesis of secondary empirical evidence, it has provided the platform for substantive studies on health workers, teachers, the police, social workers, and psychoanalysts, to name but a few. The working practices of these groups have been targeted by reforms, and they have found their professional lifeworlds transformed to a greater or lesser extent by precise demands for accountability. The critical reaction that enlists the idea of "audit society" has focused on defending the importance of professional judgment, demonstrating the decline of trust and the rise of anxiety, and analyzing the unintended consequence of auditable targets, which damage the very service quality the reforms have been intended to improve.

The auditing process is used in a variety of ways in academia, ranging from performance assessment to evaluating the effectiveness of research grants. The wide-ranging appeal of the "audit society" motif does not necessarily mean that it has been applied consistently across fields. It has often been taken as a given, supplying only a pretext for empirical analysis. In the academy, there are also variations in disciplinary sensitivities. For example, the specific methodological commitments of anthropology make it particularly vulnerable to the shorter timescales of the audit process and to the reductive nature of performance measurement. Marilyn Strathern and others have been highly critical of audit and related forms of performance measurement as they have emerged in universities in the UK. By contrast, economics, which has to a large extent internalized evaluation mechanisms as part of its own "professionalization," has been relatively silent on auditing as a specific mechanism of control and oversight.

If the concept of audit society is to be taken seriously as more than metaphor, it raises a number of serious issues for scholars. First, among these issues is the much debated boundary between audit and evaluation. Audit can be narrowly defined in the shadow of financial auditing largely as an exercise in compliance verification. From this point of view, it addresses the following question: Does this organization comply with some preestablished set of rules for the quality of its accounting? Even allowing for interpretation and discretion in the application of such rules, audit is fundamentally concerned with conformity with them. Interesting questions then concern the rules of performance themselves: How have they evolved and how are they implemented?

Evaluation, it is argued, is quite different from audit. It is an investigation of the impact of public service programs for change, assessing whether such programs succeed or fail in their own terms, and what unintended consequences they may have created. Whereas audit produces a kind of yes/no answer to compliance issues, evaluation often produces questions and ambiguity. Yet this ideal-typical distinction hides a dynamic by which the audit model expands its influence and affects evaluation in subtle forms. It is easy to see why audit is attractive to policymakers seeking simple measures of performance, and this attraction is grounded in a deeper logic that finds its purest expression in the idea of financial audit but also transcends that practice and is at work in evaluation practices as they are broadly understood. The power and impact of this logic in different institutional settings is an empirical question, but it is a logic that aspires to transform the evaluated domain according to its own operational values of visibility and traceability (the accountant's idea of the "audit trail") in the representation of organizational and individual performance.

Another criticism of the "audit society" idea is that it is a case of mistaken causality; the real driver of an "audit explosion" in the UK was in fact the heightened concern with performance improvement across a range of public services. While this is a very reasonable point to make—supported by the observed production of information that is never audited or checked in the strict sense—the concern with performance has been simultaneously a concern with the production of visible and auditable traces of performance. The logic of audit that requires individuals and organizations to be transparent and comparable plays a constitutive role in shaping conceptions of relevant performance and is at the very heart of what Michael Moran calls the attack on "club government" in the UK.

This means that "audit society" is only partly an argument about the rise of auditors, although data on increases in headcount for oversight activity are suggestive. It is more a claim about a change process that is governed by an audit logic, which in turn shapes the design of performance representations. This emphasis on process raises an important question for the scholarly understanding of "audit society" and for the attention given to the analysis of auditors as compared with those who they audit, respectively. Empirical indicators for the audit explosion can be found in the growth of audit institutions and expansion of their remit and scope. However, the less visible and more profound aspect of "audit society," which preoccupies many studies outside political science, concerns the manifest impact of the audit on professional service providers and the process by which the logic of audit is *internalized* by them.

This psychological and behavioral level of analysis reveals the conflicts and the range of strategies that individuals use to cope with new auditized demands for performance reporting. In turn, it suggests that the real "action" of the audit society happens at the level of the construction of "auditees," who are not docile subjects in Michel Foucault's sense and may put up considerable resistance to new measures of what they do but nevertheless operate in the shadow of a logic that values precision and traceability in the representation of performance. The strong form of "audit society" analysis goes further to suggest that performance representation in many different fields is shaped by this logic, with highly dysfunctional consequences.

Another methodological concern about "audit society" is that it is a case of selection on the dependent variable, a consequence of which is the failure to recognize the UK-specific nature of the phenomenon. While there is some truth in this, the argument also reveals differences between research sensitivities within and outside political science. Finding new and puzzling phenomena whose character and logic demand analysis is necessarily a precursor to the exploration of the kind of dependency and causality that provide insight into cross-sectional variation of the phenomena. In short, it is always the "dependent" variables that strike us first, generate puzzles, and provide the "value relevance" for research into their independent causes. We do not think much about this question when the dependent variable is well formed, for example, security prices in liquid capital markets, but when it is not, the investigation is more like a "depth hermeneutic" that is explicative and inductive in its movement from phenomena to their underlying logics. The audit society thesis is more than a mistaken focus on a well-formed dependent variable.

That said, cross-national variations in audit society tendencies are indeed significant, depending themselves on how fiscal and service performance issues are problematized and the sources of expertise available in different national settings and fields. The distinctive cultural position of accountants in UK society has influenced the way reform agendas have been shaped, and there is also considerable cross-sectional variation. So audit as an institutional logic with its own dynamic must compete with many others, and its hegemony in any setting is at best highly contingent.

The "audit society" idea has had wide cross-disciplinary impact because of its role in providing a label for the challenges faced by many different professional groups. However, to suggest that an audit logic has dysfunctional side effects is hardly a new one in social science. Analyses of failed dreams of mastery are well-known to political scientists and historians. The rise of audit and its underlying logic is only the latest version of this dream, and the effects, intended and unintended, are not particularly surprising. Behavioral economists have explored the "hidden costs of control" and the paradoxes of transparency. Bruno Frey and Reto Jegen have provided analyses of the "crowding" of intrinsic motivation by extrinsic efforts at control. This work overlaps with the "audit society" arguments and provides the potential for a more formal expansion of their insights.

Finally, it was suggested above that the rise of audit is closely associated with neoliberal reform agendas and the rise of the regulatory state in developed economies. And yet "audit society" is an intriguing concept precisely because it escapes easy alignment with such labels. The logic of audit embodies values of visibility and transparency that may enable greater democratic participation in service evaluation but may equally support centralist aspirations by states. Audit has been associated with a Soviet-style centralism in UK public policy, and performance audit in the Chinese context is an expression of socialist ideals. The malleability of audit and evaluation in different cultural contexts, and the capacity of their underlying logics to appear in quite different value systems, means that there is still some way to go in order to understand the complex social and organizational dynamics indentified under by the idea of "audit society." Not least among these complex dynamics is the apparent invulnerability of audit to its own failure.

Michael Power
London School of Economics
and Political Science
London, United Kingdom

See also Accountability; Auditing; Bureaucracy; Compliance; New Public Management; Performance; Policy Evaluation; Regulation; Self-Regulation

Further Readings

Cooper, A. (2001). The state of mind we're in: Social anxiety, governance and the audit society. *Psychoanalytic Studies, 3*(3–4), 349–362.

Frey, B., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys, 15*(5), 589–611.

Hood, C., James, O., Peters, G., & Scott, C. (2005). *Controlling modern government: Variety, commonality and change.* Cheltenham, UK: Edward Elgar.

Power, M. (1997). *The audit society: Rituals of verification.* Oxford, UK: Oxford University Press.

Strathern, M. (Ed.). (2000). *Audit cultures: Anthropological studies in accountability, ethics and the academy*. London: Routledge.

AUDITING

If accountability is the hallmark of democratic governance, public auditing is a means to achieve it. Auditing is a form of oversight, or examination from some point external to the system or individual in question. Technically, auditing is a form of verification by an independent body, which compares actual transactions with standard practices. Because it evaluates the relationship of what *is* against what *ought to be*, auditing is a normative operation. Public auditing is the traditional instrument to hold actors entrusted with managing public funds accountable by providing information to supervising agents, elected officials, and (sometimes) constituents about compliance with or deviations from accepted standards. Although accountability as a general aim of auditing is undisputed, unavoidable tensions with other principles of good governance arise, such as the exercise of informed discretion by elected decision makers or individual privacy rights when the corset of control is taken to an extreme.

The media, interest groups, private overseers (e.g., credit-rating agencies), international organizations, public audit institutions, and the general public increasing see modern government as an object of scrutiny. At the same time, government regulates society and delivers services through myriad national, international, nonprofit, and private organizations, making control a highly complex issue. The question is: Who oversees whom for what purpose, under whose authority, with what kind of legitimacy, and how?

The balance of autonomy and control is a classic topic for students of government. It touches on issues such as responsibility, trust, functional differentiation, authorization of execution, assurance, and accountability. At its center stands the question of how public organizations can be task efficient and accountable at the same time.

This entry is divided into four sections: The historical traditions of auditing are traced in the first section, the second one introduces different institutional arrangements of auditing, the third one sketches the main arguments of the audit explosion discourse, and the final one illustrates the politics of auditing.

Tracing the Historical Traditions of Auditing

Auditing is one of the oldest and most eminent of state functions, and early forms preceded the rise of the modern democratic state. The Latin word *auditus* means "a hearing." In ancient Rome, the organization of a "hearing of accounts" was a way in which one official would read out his accounts to another who would compare the figures. This form of oral verification was intended to prevent officials in charge of funds from misusing them. Governors of ancient civilizations, for example, Sumer, Egypt, Phoenicia, Greece, and Rome, appointed trustworthy clerks to find and punish employees for embezzlements and to protect public assets. Early Chinese writings from the 12th-century BCE illustrate an advanced understanding of the role of auditing, describing methods and practices of good economic management and the necessity of securing an independent and high-level status for the auditor. One of the earliest records of public sector financial management in India is the *Arthashastra* written by Kautilya, who was a Brahmin minister under Chandragupta Maurya, the founder of the Mauryan Empire, about 2,300 years ago. Kautilya developed bookkeeping rules to record and classify economic data, emphasized the critical role of independent periodic audits, and proposed the establishment of two important but separate offices—the treasurer and the comptroller-auditor.

The origin of public auditing lies therefore in the review of financial transaction. The concept is closely related to the development of private sector accounting techniques in 16th-century Venice, but some supreme audit institutions date their origin back to the 13th and 14th centuries. The National Audit Office in the UK sees its origin in the early 14th century, the French *Cour des Comptes* cites 1318 as the year of its inception, professional forms of supervision and control of resources in Mexico date back to the Aztec Empire. These audit institutions were not independent of the ruler or crown and were designed to control the financial conduct of the crown servants. Their aim was to protect the interests, property, and wealth of a ruler at a time when the state was still considered private property. In concert with trends of the 18th and 19th centuries—the diminishing role of family and personal loyalty, the increasing division of labor in economy and state (functional differentiation), the emergence of big government, new ideas about responsible behavior of public servants—audit institutions changed in scope, quality, and quantity.

As parliament gained power to grant the financial means to execute public tasks, the need arose to oversee the respective administrations or institutions using the financial resources. The objective was to ensure that the regularity and probity of the expenditures was in accordance with the defined aims of parliament. The state budget was and is the prerequisite for this form of formal legal rationality. Its verification rests on checking receipts against authorized state budget expenditures and vice versa.

The clearest transformation was from internal to external auditing and with it the emergence of independent audit institutions, often considered to be the extended arm of the parliament via the executive government. Over time, they gained a constitutional status in many countries, which secured their statutory and operational independence from executive interference. In many countries, however, especially those under authoritarian rule, this was the exception rather than the rule. In the former German Democratic Republic, an independent audit system did not exist to execute the directives of the Ministry of Finance, which made the Central Committee of the Communist Party a political tool to oversee the completion of economic plans in a state-directed economy. The attribute of constitutionally secured independence of the auditor from those holding and executing power (relational distance), for example, cabinet, bureaucracy, parliament, and the auditor's role in ensuring good governance have become, as many see it, the key to the auditor's legitimacy in democracies. The many authoritarian regimes that lack an autonomous body of oversight in effect allow the instrumentalization of spending for partisan purposes. The *Mexico Declaration on SAI Independence* of the International Organization of Supreme Audit Institutions advocates principles of good auditing and tries to establish common standards of auditing among its global member organizations. It specifies organizational, functional, and financial independence as the core prerequisite to accomplishing tasks objectively, effectively, and without external interference.

The state audit evolved from a rather crude tool to prevent chaos and ruin in public finance to an advanced instrument for safeguarding responsive, accountable, open, regular, and efficient government. *Compliance auditing* (verification of accounts, legality of expenditures, the organization's adherence to regulatory guidelines) and *financial auditing* (the overseeing of public accounts, assessing accuracy and fairness of an organization's financial statements) were the primary purposes of auditing; *performance auditing* (examination of how government programs operate and how well they achieve the defined objectives) came somewhat later, gaining an important place in the auditing realm with the new public management reforms. The latter is often referred to as *value-for-money auditing*, or evaluating the efficiency, effectiveness, and economy of programs and policies. Such a focus is supposed to give way to comparisons with performance criteria of best practices from the private sector. Its critics claim that performance auditing has contributed to an overstretched meaning of auditing: For example, conceptual boundaries with forms of evaluation have started to blur.

Auditing as an Institutional Arrangement

Internal versus external auditing is the most basic differentiation of types of auditing. It refers to the organizational status of the auditing entity relative to the one it audits. *Internal* means that the auditor is part of the organization and the organization's employee; *external* refers to an external oversight body or person operating independently of the entity being audited. Because internal auditors are part of the organization under scrutiny, they work for the benefit of the management of the organization under examination and their insights remain mostly disclosed only within the management. In contrast, external auditors review financial transactions and/or statements and serve third parties, for example, parliament, president, who need reliable financial information about those entrusted with the management of public funds.

Models of Auditing: State, Private, and People's

The audit of public sector organizations staffed with civil servants differs. The most general classification is based on the question of who conducts the audit. Here we can distinguish between state, private, and people's auditing. *State auditing* can be divided into three subgroups according to government level:

1. supreme audit institutions (e.g., the British National Audit Office) examine the affairs of federal government departments;
2. second- or third-level audit institutions (which can be fully independent of other government levels) exist in states, provinces, and municipalities of federal systems to audit their own government entities; and
3. internal auditing bodies operate within any governmental department or agency.

Professional accounting firms charged with *private auditing* conduct special-purpose audits (such as reviewing public enterprise accounts) and have increased in number and scope in recent decades along with the growth of government-owned business corporations in the public sector. The third audit model, *people's auditing*, was exercised in the former Soviet Union by the *Narodnyi Control* (Soviet People's Control) as part of a combination of professional auditors with millions of volunteers as so-called people's inspectors who conducted examinations and inspections in their spare time. This system was meant to educate the masses and be a training ground in the transformation of the country into a communist society. According to the communist ideology, the professional state apparatus would eventually wither away and the ordinary people would take over.

Organization of Auditing: Court or Administrative System

State (supreme) audit institutions can be classified into two broad groups according to their organization: the *cameral* or *court system* and the *administrative* or *monocratic system.*

The Cameral System

In the cameral system, the audit institution is headed by a group of auditors whose independence is ensured by their judicial status and who examine governmental accounts as well as the uses of public funds. To this ideal-type model belong most of the Roman law–based countries, such as not only the Corte dei Conti in Italy, the French Cour des Comptes, and the Tribunal de Cuentas in Spain but also the European Court of Auditors, the German Bundesrechnungshof, the Dutch Algemene Rekenkamer, and the Auditoría General de la Nación of Argentina in Latin America. In Italy, France, and Spain, the auditor often acts as an administrative tribunal with quasi-judicial powers in administrative matters, whereas the variant of this collegiate system in Sweden, the Netherlands, and Germany is without jurisdictional authority. The strong legalistic foundation in both variants of the "cameral system" privileges legal and financial compliance over performance auditing. Though in principle a collegiate model in which everyone enjoys the same rights and duties, the institution is commonly headed by a kind of *primus inter pares*. As a "first among equals," he or she can exert, as in the German case, considerable influence by selecting the heads of sections and tailoring the areas of operation. Despite these far-reaching organizational competencies, the head of the institution cannot interfere with actual matters of auditing or the content of inspectors' decisions.

The Administrative or Monocratic System

In the United Kingdom (UK), Canada, and most of the former British Empire dominions and colonies, as well as countries such as Mexico, Chile, and Colombia, the state audit institution is organized under the administrative or monocratic system, which is similar to a hierarchically structured government department—that is, headed by a single person called the comptroller or auditor general and acting as an auxiliary institution to the legislature, although charged with ample autonomy. The head of the institution enjoys an independent status, but the institution itself functions according to hierarchical principles of lines of command. Under this system, controls are intended to correct rather than penalize. Supreme audit institutions in these countries focus more on *ex post* auditing, rather than *ex ante* control, and emphasize financial and performance auditing over compliance control.

Interrelating With Others: Parliamentary-Led or Executive-Led Auditing

Audit institutions are embedded in political environments, which is exhibited by the way they interrelate with other actors. Though they all enjoy statutory independence, differences arise with regard to instrumental institutions: the one to which they report, the one that appoints the head of the Supreme Audit Institution, and the one that assigns the annual budget. Despite their statutory independence and regulations to protect that status, auditors are dependent on, for example, national legislative bodies or governmental budget support—forms of leverage that can be used to compromise the auditor's independence.

In Argentina, Canada, Germany, the United States, Israel, Mexico, and the UK, the parliament or a special committee appoints the Auditor General, acts as supervisor, and decides on the annual budget. This system has been labeled *parliamentary led,* as the auditor is considered the extended arm of the legislator. In an *executive-led* system such as France (also Ecuador), the "Premier Président" and other magistrates of the Cour de Compte are appointed by a decree from the President of the Republic. Appropriations are officially related to the Prime Minister but have special authorization by Parliament. Reports are directed to the president and legislative branch of government. China also has an executive-led arrangement: The auditor reports to the Premier of the State Council (the head of the Chinese government) and the ministry of finance controls the budget. Sweden's Riksrevisionen, which until 2000 was similar to the executive-led model, now carries out its state audit functions under the auspices of Parliament.

Auditing as a "Neurosis"

Balancing independence with control is a classic topic in public sector research. The new public management (NPM) hype sparked an increased interest in performance auditing, primarily understood as an assessment of efficiency goals and other indicators of economic success with a corresponding system of controlling and reporting. At its center stands the question of how public organizations can be task efficient and accountable while being relatively free of bureaucratic impediment.

In the past 2 decades, almost every local government and many national administrations in OECD countries have implemented some kind of NPM reform. This "wave of modernization" has in many cases changed the way governments and local authorities operate and eventually also the audit regime. Although the NPM doctrine does not prescribe a well-defined list of reform steps, the stereotypical tool box includes privatization, deregulation, outsourcing, service delivery competition, semiautonomous service centers, and—particularly in central and federal government reform—executive agencies, performance standards and measures, increased consumer power, and employment of professional managers and noncareer personnel. Auditing bodies, it is claimed, respond to NPM-type reforms by institutional, procedural, and cultural acclimation. Increasing coproduction of public goods and services by public, private, and nonprofit organizations seems to pose additional challenges to systems of public auditing that are supposed to counter centrifugal tendencies of decentralized agencies and public enterprises.

The conceptual overstretching of auditing is often associated with a debate on the rise of the so-called audit society. An increasing public and academic concern about unethical behavior of public officials, an evaluation-obsessed political and social context, and the observation that modern Western societies would evolve into low-trust societies led observers to conclude that the "audit society" is a way of dealing with this type of neurosis. This diagnosis had its origin in Michael Power's description of a society obsessed with overseeing almost every domain of public activity and the paradoxical effect that it does *not* lead to an increase in trust in government. This audit mentality is very often decoupled from its real effects on public sector accountability and has become an end in itself. Auditing is not seen as a rational reaction to the well-defined problem of accountability; rather, the very concept creates, constitutes, and shapes the public perception for which it is a solution. Therefore, by "audit explosion," Power refers to the growth of audit and related monitoring practices associated with public management reform processes. As a multifaceted concept, the explosion is not only an increment of auditing activities in quantitative terms but also a shift in qualitative perspective, mainly toward the so-called value-for-money audit and other internal control systems. It also means that oversight bodies offshoot into many different fields such as health care and education. The concept of an audit society has been criticized for being too narrowly attributed to the UK or other English-speaking countries, such as Australia or New Zealand, meaning that it does not sufficiently capture the reality of developments in other countries, for example, the United States or Germany.

The Politics of Auditing

Audit institutions do not commonly have the power to set rules or sanction the deviant behavior of auditees. In principle, they lack the capacity to exert direct hierarchical control. This is not undisputed, as the Audit Office of Finland, the Office of the Comptroller General of Chile, and the Court of Audit in Spain are endowed with competencies to prosecute. The Italian Corte dei Conti and the French Cour de Compte also enjoy quasi-judicial powers in administrative matters. Despite these examples, the foundation of their authority in general, however, rests on their constitutionally guaranteed independence from political interference, their professional expertise, the skill to informally cooperate with auditees in order to improve processes, and the ability to ally themselves with like-minded organizations (such as the public accounts committee in parliament) to exercise indirect control. Providing information to other institutions about a specific body's financial management is the dominant way of exerting influence. The naming and shaming of deviant behavior through the publication of examination results must be seen as *ultimo ratio*. On one hand, publicizing mismanagement is a powerful lever; on the other, it politicizes their alleged independence as the auditor is drawn into the struggle between political competitors. For this reason, the evaluation of government policies by auditors is a critical matter.

Auditing is per se political in nature. It operates in the nexus of politics and policies. The tension between politicization and institutional independence is permanently felt in the arena of political competition for several reasons. First, the broad legal mandate with the discretion to look into almost any government activity and its potential to unveil, for example, untoward spending practices unwittingly makes the auditor a player within the political game. Second, the nature of the auditing workforce—empowered, well positioned in government, specifically skilled—impels political parties to strategically place favorites. Third, an independent producer of creditable and presumably apolitical information is a highly desirable partner in political arguments. Auditing as a political power or blaming tool is most evident when it is used to detect and deter deviant behavior, such as capitalistic tendencies in communist countries, or to achieve certain policy or partisan objectives. The claim of avoiding policy issues and focusing on the operation of the apparatus is therefore understandable, albeit next to impossible, especially in the context of issues framed not objectively but within the perceptions of political actors.

The principle of the auditors' independence is definitely a necessary condition for a well-functioning system. Other conditions are more general traits of the political system, such the audit courts' dependence on the democratic culture of a country, the respect for democratic institutions, the belief in a division of powers, and respectful behavior among those holding power. The main challenge in the development of auditing in modern states is therefore the acceptance of the complementarities of the auditor's role by relative political actors and the positive use of its potential to improve government operations, ensure good public records, and increase public trust.

Patrick von Maravic
Zeppelin University
Friedrichshafen, Germany

See also Accountability; Budgeting, Rational Models; Compliance; New Public Management; Responsibility

Further Readings

Barzelay, M. (1997). Central audit institutions and performance audits: A comparative analysis of organizational strategies in the OECD. *Governance, 10*(3), 235–260.

Chelimsky, E. (1985). Comparing and contrasting auditing and evaluation: Some notes on their relationship. *Evaluation Review*, *9*(4), 483–503.

Funnell, W. (1994). Independence and the state auditor in Britain: A constitutional keystone or a case of reified imagery? *ABACUS*, *30*(2), 175–195.

Geist, B. (1981). *State audit: Developments in public accountability*. London: Macmillan.

Gong, T. (2009). Audit for accountability in China: An incomplete mission. *The Australian Journal of Public Administration, 68*(Suppl. 1), S5–S16.

Grasso, P. G., & Sharkansky, I. (2001). The auditing of public policy and the politics of auditing: The U.S. GAO and Israel's State Comptroller. *Governance*, *14*(1), 1–21.

Hood, C. (2004). Controlling public services and government: Towards a cross-national perspective. In *Controlling modern government* (pp. 3–24). Cheltenham, UK: Edward Elgar.

Pollitt, C., Lonsdale, J., & Girre, X. (1999). *Performance or compliance? Performance audit and public management in five countries*. Oxford, UK: Oxford University Press.

Power, M. (1997). *The audit society: Rituals of verification*. Oxford, UK: Oxford University Press.

Santiso, C. (2009). *The political economy of government auditing*. New York: Routledge.

Scott, C. (2003). Speaking softly without big sticks: Meta-regulation and public sector audit. *Law & Policy*, *25*(3), 203–219.

Sutherland, S. L. (1986). The politics of audit: The federal office of the auditor general in comparative perspective. *Canadian Public Administration*, *29*(1), 118–148.

AUTHORITARIAN REGIMES

Until the early 1960s, there was a fundamental dichotomy in political science between democratic and totalitarian regimes. Authoritarian regimes were often treated as a type of totalitarian regime. However, with the collapse (Germany) and the subsequent transformation (in the former USSR) of totalitarian regimes after World War II and in the 1950s, authoritarian regimes have become the dominant form of nondemocratic government.

This entry provides a general definition of authoritarian regime as well as exploring subtypes. Among the many subtypes of authoritarian regimes, the currently most widespread one is the "electoral authoritarian" regime that combines a facade of multiparty elections with a variety of authoritarian controls that, in effect, deprive such elections of any democratic significance. Therefore, this entry, while briefly discussing other subtypes of authoritarian regimes, will particularly focus on electoral authoritarianism. Because such regimes are the dominant regime type in the Middle East and North Africa, a special emphasis will be put on this region.

A Theory of Authoritarian Regimes

Juan J. Linz developed the first full-fledged theory of authoritarian regimes, as distinct from both democracies and totalitarian regimes, as well as from many other types of other nondemocratic regimes (traditional autocracies, sultanistic regimes, semiconstitutional monarchies, oligarchic democracies). As early as 1964, Linz defined authoritarian regimes on the basis of three main characteristics: (1) limited pluralism, (2) the presence of a mentality rather than an ideology, and (3) the absence or low level of political mobilization. He developed and refined his theory in his later work (especially *Totalitarian and Authoritarian Regimes*, 2000). More recently, such regimes have started to attract greater scholarly attention for several reasons. First, the emergence of military–bureaucratic authoritarian regimes in much of Latin America and elsewhere as a new type of authoritarianism required more detailed research into their dynamics. Second, after the global resurgence of democracy or the "third wave of democratization" in the 1980s and the 1990s, there emerged a still newer type of authoritarianism generally called "electoral authoritarianism." This type lies in the gray area between liberal democracies and "closed" authoritarian regimes, an area that they share with "electoral democracies." Hence, in this entry, while all subtypes of authoritarian regimes will be discussed, the emphasis will be on these intermediary types.

Limited Pluralism

Starting with the three-dimensional scheme of Linz, limited pluralism contrasts with the almost unlimited pluralism of liberal democracies and the almost total absence of pluralism in totalitarian regimes. The degree of limited pluralism varies among different subtypes of authoritarian regimes. They may exclude only political opposition groups or may exclude certain types of civil society institutions as well. Such exclusion may be based on legal regulations or de facto restrictions. As this observation indicates, the degree of limited pluralism is a critical variable in distinguishing the subtypes of authoritarian regimes, as will be spelled out below.

Mentality Versus Ideology

A second defining element of authoritarian regimes is the presence of a "mentality" rather than an "ideology." In Linz's view, in contrast to the more rigid, systematic, rationally developed, utopian, and comprehensive nature of ideologies, mentalities are less systematic and more amorphous ways of thinking and feeling, more emotional than rational. Although the distinction is not clear-cut and there is a large gray area in between, "mentalities," as recalled by the German sociologist Theodore Geiger, are, actually, more or less ambiguously developed values, such as fatherland, nation, order, hierarchy, tradition, race, ethnicity, and others. The key political characteristic of them is that, precisely because they are ambiguous and intellectually underdeveloped, creating broader political coalitions around them to build the bases of regime legitimacy is much easier than with precise, developed ideologies (Linz, 2000).

Lack of Political Mobilization

The third element is the absence of extensive political mobilization, again a crucial difference with totalitarian regimes, even though authoritarian regimes may also have followed mobilizational policies at some point of their development. Actually, this variable is closely related to the second one, since in the absence of an ideology, it is difficult to pursue effective mobilizational policies and to attract the young, students, and intellectuals.

Borderline Cases

Admittedly, in terms of all three criteria, the borderline separating authoritarian regimes from other types of nondemocratic regimes on the one hand and from pseudodemocracies or semidemocracies on the other is a rather blurred one. Thus, an authoritarian regime that allows a fairly large area of limited pluralism in nonpolitical fields and even tolerates and attempts to co-opt some of its political opponents, maintains a fairly flexible mentality that leaves room for different interpretations, and shows little interest in mobilizing the masses in a certain ideological direction may be said to approach a pseudodemocracy, while there may be cases where it is difficult to distinguish authoritarianism from totalitarianism. Hence, this large gray area poses a number of difficult borderline problems, and it requires an effort to distinguish subtypes of authoritarian regimes. Such difficulties of classification are further aggravated by the fact that many regimes combine elements from different types and stress different elements in different phases of their history. Authoritarian regimes are often complex systems combining elements of contradictory models in uneasy coexistence.

Subtypes of Authoritarian Regimes

Linz distinguished the following subtypes of authoritarian regimes using his two criteria of the degree of limited pluralism and the apathy mobilization axis:

- bureaucratic-military authoritarian regimes,
- organic statism,
- mobilizational authoritarian regimes in postdemocratic societies,
- postindependence, mobilizational authoritarian regimes,
- racial and ethnic democracies, and
- posttotalitarian authoritarian regimes.

Although space limitations do not allow for a detailed analysis of each of these subtypes, it will be useful to discuss whether this typology is as meaningful today, many decades after Linz's masterful essay.

Certainly, the collapse of the Soviet type of totalitarianism has made the fundamental distinction between totalitarian and authoritarian regimes much less meaningful at least for practical purposes. Probably the only Soviet-type totalitarian regime today is that of North Korea. Linz classifies China and Cuba as posttotalitarian. Certain former Soviet republics of Central Asia seem to combine posttotalitarian and sultanistic elements. On the other hand, Central and Eastern European countries that were part of the former Soviet bloc seem to have made a successful transition to democracy, while some of the former Soviet republics (Russia, Belarus, Ukraine, Moldova, Transcaucasian republics, Kyrgyzstan) are still experiencing serious problems of democratic consolidation, probably because of their much longer totalitarian past and difficulties associated with the problem of "stateness."

Another interesting borderline case between authoritarianism and totalitarianism is Iran, where an almost totalitarian interpretation of a religious ideology is combined with elements of limited pluralism. Under the Islamist regime, Islam has been transformed into a political ideology with a totalitarian bent, and the limited pluralism is allowed only among political groups loyal to the Islamic revolution. Furthermore, the electoral and legislative processes are under the firm control of tutelary institutions such as the Supreme Leader (appointed by the Assembly of Experts composed of theologians), the Council of Guardians with veto powers over all laws and all candidates seeking elective office, and the Expediency Council added to the Constitution in 1990 as the final arbiter in cases of disagreement between the Majlis (parliament) and the Council of Guardians, all inspired by Khomeini's doctrine of "the supremacy of the Islamic jurisprudent" (*Velayet-e Faqih*). These tutelary institutions effectively blocked the moderate reform attempts of President Mohammad Khatami (elected twice in 1997 and 2001, himself a cleric loyal to the Islamic Revolution), even though he had the support of the Majlis elected in 2000 and dominated by the reformists. Thus, Iran's limited pluralism appears very limited indeed.

The momentous socioeconomic and political changes in the past 3 or 4 decades, at the top of which stands "the third wave of democratization," made some of Linz's subtypes less relevant today than they used to be. Thus, with the transition to democracy in the South African Republic, the leading example of so-called racial or ethnic democracies disappeared. With the demise of totalitarianism, defective and pretotalitarian regimes are no longer a contemporary category. With the successful transition to democracy in former Soviet bloc countries of Central and Eastern Europe, posttotalitarianism is limited to a few countries such as China and Cuba, as was alluded to above. Mobilizational authoritarian regimes in postdemocratic societies also seem to be a category of mainly historical significance. Such regimes emerged in the past in countries with some democratic experience and a degree of political mobilization. Therefore, the authoritarian regimes had to use a mass organizational party and other mass organizations in order to exclude the defeated social force from political participation. However, with the worldwide preeminence of democratic ideology at present, it seems exceedingly difficult to establish such a regime in a country with some prior experience in democratic politics and mass mobilization.

Similar observations can be made about the subtype called "organic statism" by Linz. Such regimes attempt to combine the Catholic corporativist social doctrine with fascist elements but are distinct from fascist-totalitarian regimes. In these regimes, often described as coorporatism or organic democracy, a variety of social groups or institutions that are created or licensed by the state are allowed to participate in some degree in the political process. Linz admits, however, that this model, popular in the 1930s, lost its attraction for a number of reasons. Probably, the most important reason was the disengagement of the Catholic Church from its commitment to essentially undemocratic organic theories of society. A number of "populist regimes" that emerged especially in Latin America in the 1930s and later, such as those of Getulio Vargas in Brazil and Juan Peron in Argentina, interestingly, combined elements of corporatism with those of "electoral authoritarianism" to be discussed later. These regimes displayed a strong developmentalist, nationalist, statist, and elitist character. With the failure of their import substitution–based development strategy and the emergence of the currently dominant free market economy, however, such regimes seem to have lost their appeal.

Finally, the usefulness of another subtype called "tutelary democracies" by Edward Shils is also open to doubt. Recent research has suggested, for example, that Kemalist Turkey, usually presented as the prototype of such regimes, in fact bore closer resemblance to corporatist or organic statist authoritarian regimes. Indeed, it is difficult to conceive of an "enlightened" leader or a group of leaders dedicated to "educate" their people in democratic values and prepared to lose power as the ultimate outcome of the democratization process.

The remaining subtypes of authoritarian regimes with some contemporary relevance seem to be the military–bureaucratic authoritarian regimes and the postindependence mobilizational single-party regimes. The latter are normally associated with a war or struggle of independence from colonial rule. Countries in Black Africa and the Maghreb provide examples of this type. Linz observes that these regimes emerged in countries of low economic development where the colonial rulers had not allowed the growth of an indigenous professional or bureaucratic middle class but destroyed or weakened traditional precolonial authorities. Thus, the new nationalist leadership, mostly trained abroad, successfully mobilized the grievances of the native population against the foreign rulers. As expected, the desire for independence obscured the importance of other cleavages. After a short period of experience with democratic constitutions inherited or inspired by their former colonial rulers, most of these regimes were transformed into single-party regimes, the most notable examples of which were Algeria, Tunisia, Ghana, Ivory Coast, Senegal, and Tanzania. The single-party regimes claimed to represent the entire nation and defined their chief mission as nation building. However, their mobilizational and ideological components decayed rapidly. The role of the mass mobilizational single-party regime lost its preeminence. Consequently, they moved in the direction of either military–bureaucratic or organic statist regimes. This process can be compared with "detotalitarianization by decay," namely, the loss of the revolutionary ideological élan and of the mobilizational capability of the regime, which also explains the sudden collapse of the Soviet Union probably better than anything else.

Turkey was one of the earliest examples of single-party regimes and, as such, attracted the attention of a great number of Western scholars, including Maurice Duverger. The Republican People's Party (RPP) was founded by Kemal Atatürk, the leader of the War of Independence (1919–1922) in 1923, and it consolidated its monopoly in 1925 by outlawing the only opposition party, the Progressive Republican Party. It maintained such monopoly until the end of the World War II, when the party leadership allowed the formation of an opposition party, the Democratic Party and lost power to it in the first free election of 1950. The RPP (and its ancestor, the Society of Defense of Rights for Anatolia and Rumelia) was successful in achieving a significant degree of popular mobilization during the War of Independence, appealing to and co-opting Muslim religious orders, Kurdish tribes, radical republicans, and loyal monarchists, and so on. After the victory, however, the RPP was quickly transformed into an "exclusionary" single party, in the words of Samuel Huntington. It showed little interest in mass mobilization and concentrated its efforts on nation building and secularizing the state and society by top-down authoritarian methods, almost a "mission impossible" in a deeply divided, pluralistic, and devoutly religious society.

The military–bureaucratic authoritarian regimes emerged as probably the most widespread type of authoritarianism in the 1960s, especially in Latin America but also elsewhere. Guillermo O'Donnell and others have argued that the emergence of such regimes as a new type of authoritarian rule in Latin America was linked to the failure of import substitution–based industrialization (ISI). The early phases of this industrialization strategy gave rise to populist coalitions that brought together national industrialists, urban professionals, and the emerging urban working class. Both components of the populist coalition benefited from this expansionist policy: Import-substituting industrialists were heavily protected against foreign competition by import quotas, tariffs, and foreign exchange regulations; the urban working class, in turn, obtained employment opportunities, union rights, welfare benefits, and relatively higher wages because the growth of the domestic market was supposed to further stimulate the ISI.

The early, easy phase of the ISI, however, was soon exhausted. Once the limits of growth of domestic markets had been reached, the inefficient, internationally noncompetitive national industries failed to produce the necessary foreign exchange earnings. Thus, ISI-based economies faced severe problems, such as balance-of-payment deficits, foreign exchange shortages, low or negative growth rates, inflation, and unemployment. The economic crisis, in turn, led to the radicalization of the popular sector (the industrial working class) and to increased political polarization. This created fears among the urban middle classes of a Cuban-style radical leftist revolution. With the support of this middle class, the military took over in many important countries of Latin America such as Brazil, Argentina, Chile, and Uruguay and established what is known as the bureaucratic–military authoritarian regimes. The fact that these regimes emerged in the most highly developed countries of Latin America lends support to O'Donnell's economy-based explanation.

Such regimes differ from earlier types of military regimes in Latin America (*poder moderador*) in that the latter were of relatively short duration and intended to solve a particular constitutional crisis, while the former were of much longer duration and attempted to change the economic structure in a more free market, export-oriented direction. To accomplish this objective, they tried to demobilize the already activated popular sector by severely restricting its political participation opportunities. In this effort, they also enlisted the help of a large group of experts, technocrats, and managers with problem-solving capabilities. Hence, these regimes are often called "military–technocratic–bureaucratic authoritarian regimes" to distinguish them from the more bureaucratic–military–oligarchic authoritarian regimes of interwar Eastern Europe.

Although the bureaucratic authoritarian regimes have been analyzed mostly within the Latin American context, the applicability of this model is not limited to that continent. In fact, the National Security Council (NSC) regime in Turkey (1980–1983) bears strong resemblances to its Latin American counterparts. It also followed a period of deep economic crisis as a result of the failure of an ISI strategy, accompanied by intense political polarization, radicalization of the working class, and the increasing fears of the middle classes. Like its Latin American counterparts, the NSC regime pursued policies aimed at shifting the basis of the economy from an ISI to an export promotion model, while trying to demobilize the working class and depoliticize the society at large.

Parallel to the "global resurgence of democracy," pure and simple military regimes have virtually disappeared, except perhaps for a transitional period. As Larry Diamond observes, the interventionist armies today try to either legitimize their rule by running for president in contested but manipulated elections or carve out for themselves a large autonomous sphere in the form of tutelary powers and reserved domains behind a veil of civilian, multiparty rule.

Electoral Authoritarianism

While some of the subtypes in Linz's original typology have lost their contemporary relevance, new borderline problems have arisen as a result of the rise of democratic ideology as the almost unrivaled source of legitimacy. Many otherwise quite authoritarian regimes have been obliged to accept at least some outward trappings of democratic regimes. As a result, while in the past, the boundary between authoritarian and totalitarian regimes was not very clear, as opposed to the relatively well-defined boundary between nondemocratic and democratic systems, today the reverse seems to be true. Indeed, there seems to be a large gray area between authoritarian regimes and certain types of democracies variously called "electoral democracies," "semidemocracies," "pseudodemocracies," and so on. A newly coined term, *electoral authoritarianism*, suggests the difficulty of arriving at a neat classification. Electoral or competitive authoritarian regimes are not among the subtypes in Linz's original typology, for understandable reasons. As Diamond argues, this type of hybrid regime is very much a product of the contemporary world. The worldwide trend toward democracy has forced many authoritarian regimes to mimic some forms of electoral competition. Andreas Schedler argues, in the same vein, that on the bases of 2001 data, out of 151 countries in the "developing world," 36 (23.8%) are liberal democracies, 32 (21.2%) are electoral democracies, 58 (38.4%) are electoral authoritarian regimes, and 25 (16.5%) are closed authoritarian regimes. Thus, electoral authoritarian regimes make up more than two thirds of all autocracies.

The breadth of this gray area between liberal democracies and full-scale (closed) authoritarian regimes may necessitate the search for further subdivisions within this category. Thus, Schedler distinguishes between electoral democracies and electoral authoritarianism, while Diamond, in a more nuanced distinction, divides it into four subtypes as electoral democracies, ambiguous regimes, competitive authoritarian regimes, and hegemonic electoral authoritarian regimes. Both authors agree that the boundaries between these regime types are "blurry" or "foggy" and that it is difficult to fit elusive realities into ideal types. The distinction between liberal and electoral democracies is based on variables other than electoral competitiveness, since both conduct reasonably free and competitive elections as a minimum condition for democracy. Electoral democracies fail to institutionalize the other vital dimensions of liberal democracies, such as the rule of law, political accountability, full protection of civil and political rights, and so on. Another feature distinguishing electoral democracies from liberal democracies may be the presence of tutelary powers and reserved domains in some of them outside the purview of elected officials, most often in favor of the military but also sometimes in favor of clerical leaders as in the case of Iran. Turkey, where free and fair elections have been regularly held since 1950, but the military have enjoyed significant tutelary powers and reserved domains since its intervention in 1960, and even more so after its intervention in 1980, is a good example of such cases.

On the other hand, the difference between electoral democracies and electoral authoritarian regimes centers on the fairness and competitiveness of the electoral process. The criteria for free and fair elections are well known and need not be repeated here. However, as Schedler puts it, historically, elections have been an instrument of authoritarian control as well as a means of democratic governance. And at present, there are many regimes where multiparty elections are regularly held, but many basic democratic norms are systematically violated. Therefore, they should be classified as authoritarian regimes. Indeed, the authoritarian use of elections is nothing new, as attested by an old but influential book, *Elections Without Choice* (1978), edited by Guy Hermet, Richard Rose, and Alain Rouquié, providing earlier examples of such regimes. As Schedler rightly observes, "the menu of manipulation" available to authoritarian rulers is quite rich and varied, ranging from banning parties and disqualifying candidates to repressive policies against dissenters during election campaigns, fraud and intimidation, informal disenfranchisement, unequal use of state resources, adoption of favorable redistributive rules, and so on. Certainly, one can further subdivide the electoral authoritarianism category into competitive authoritarian and hegemonic authoritarian subtypes, as did Diamond, depending on their degree of competitiveness. In his view, competitive authoritarian regimes are defined by the presence of a significant parliamentary opposition, whereas in the hegemonic subtype, elections are largely an authoritarian facade: The ruling or dominant party wins almost all the seats. Similarly, in hegemonic electoral authoritarian regimes, the president is elected with a great majority of the popular vote. Diamond admits, however, the difficulty of distinguishing between the two types. At any rate, the difference seems to be one of degree rather than of kind. No doubt, they both firmly belong to the authoritarian type, but both differ from full-fledged or "closed" authoritarian regimes.

The blurry borderline between electoral democracies and electoral authoritarian regimes poses difficult methodological problems. As Schedler rightly argues, under conditions of electoral manipulation, official election figures do not faithfully reflect the actual distribution of citizen preferences. However, it is difficult to observe electoral manipulations, since these are usually "undercover activities," constituting "an impenetrable black box." Nevertheless, available information on the state of political rights in general may give us clues about the extent of electoral manipulation. Thus, unequal access to state resources or mass media, restrictions on freedom of expression and association, bans on political party activities, adoption of electoral rules favorable to the governing party, gerrymandering, and so on may provide reasonably reliable indicators to recognize an electoral authoritarian regime. Observations by international election observers, if available, may provide another source of information.

As was pointed out above, electoral authoritarian regimes constitute the largest group of countries in the developing world. Certainly, the rise of this type is closely associated with the global resurgence of democracy. With the collapse of Soviet-type totalitarianism, democracy has become the only legitimate form of government in the eyes of a much greater portion of the world's population. As a result of a favorable international climate, active promotion and encouragement of democratic change by major Western powers, the desire to gain international respectability, and domestic pressures associated with rising levels of socioeconomic development, some forms of authoritarianism have become less sustainable. Thus, most hybrid regimes today are quite deliberately *pseudodemocratic*, dreaming, in Schedler's words, of reaping the fruits of electoral legitimacy without running the risks of democratic uncertainty. They mimic democratic institutions not only in the electoral field but also by creating mimic constitutional courts and adopting other outward trappings of a system based on the rule of law. Conceivably, such institutions create potential areas of dissidence and conflict, even if they remain far from ensuring democratic accountability. Nevertheless, the concept of electoral authoritarianism centers on the electoral area, as its name indicates.

Thus, electoral authoritarian regimes combine two contradictory dynamics in uneasy coexistence. On the one hand, as Schedler points out, by permitting multiparty, competitive elections, they recognize the principle of democratic legitimacy instead of other sources of legitimacy (such as revolutionary, transcendental, communitarian, charismatic, or substantive). On the other hand, they subvert it in practice using various instruments of authoritarian manipulation. As opposed to democratic regimes in which the rules of electoral game are accepted and respected by all parties, in electoral authoritarian regimes, competition is not only "within rules" but also, and more important, "over rules." On the other hand, as opposed to closed authoritarian regimes, the rulers have to find a balance between "electoral persuasion" and "electoral manipulation."

This mixture of contradictory elements makes the future of electoral authoritarian regimes highly uncertain and unpredictable. Again, as Schedler points out, the game of authoritarian elections may lead to a process of gradual democratization as in Mexico and Senegal. It may lead to democracy through the sudden collapse of authoritarianism as in Peru and Serbia. It may also end up in an authoritarian regression. Or such regimes may survive for an extended period, through various successful survival strategies of the rulers.

At the other end of the democracy–autocracy continuum, "closed authoritarianism" seems to be the relatively smaller category, but a stable one. Thus, during the past 10 years, such a nonfree regime went from 25% of all states (48 in total; *Freedom in the World 2000–2001*) to 24% (47 in total; *Freedom in the World 2009*). They may combine some elements of "sultanistic regimes," considered a separate type by Linz. He has argued that the exceedingly personalistic and arbitrary nature of such regimes distinguishes them from the more institutionalized and predictable behavior of the authoritarian regimes. In sultanistic regimes, the ruler exercises his power unrestrained by legal rules or ideological commitments. Public bureaucracy becomes the personal entourage and servants of the ruler. Classic examples of sultanistic regimes are the Trujillo regime in the Dominican Republic, the Batista regime in Cuba, the Somoza regime in Nicaragua, the Duvalier regime in Haiti, the Pahlavi regime in Iran, and the Marcos regime in the Philippines. More recent examples may include the Ceauşescu regime in Romania, the Saddam Hussein regime in Iraq, and the North Korean regime.

If we look at the regional distribution of authoritarian regimes, we see its highest incidence in the Middle East and the North Africa and its lowest in Eastern Europe, with sub-Saharan Africa in between. According to Schedler's figures (as of early 2006), 11 sub-Saharan African countries (Burkina Faso, Cameroon, Chad, Ethiopia, Gabon, Gambia, Guinea, Mauritania, Tanzania, Togo, and Zambia) are in the electoral authoritarian category. According to his figures based on 2001 data, 52.6% of the Middle Eastern and North African regimes are electoral authoritarian regimes and 42.1% belong to closed authoritarianism category. Israel is the only liberal democracy in his classification, with Turkey included in the electoral authoritarian category. In contrast, 57.9% of Eastern European states are liberal democracies, 15.8% electoral democracies, 26.3% electoral authoritarian regimes; no Eastern European regime is in the closed authoritarianism category. The success of the Eastern European cases can be explained by the geographical and cultural proximity of Western democracies and many forms of international influences, including the European Union conditionality. The Middle East and North Africa, on the other hand, stand as the major geographical region of the world least affected by the third wave of democratization, a fact that requires a detailed explanation.

Islam and Democracy

One explanation for the limited democratization in the Middle East and North Africa, based on the alleged inherent incompatibility between democracy and Islam, the dominant religion of the region, must be treated with utmost care. It has been argued by Orientalists, for example, that many notions associated with Western democracy, such as the notions of popular sovereignty, representation, elections, secular laws, an independent judiciary, and a civil society composed of a multitude of autonomous groups, are alien to the Muslim political tradition. Such statements abound in the Orientalist literature and have deeply influenced the Western conceptions of Islam, so much so that Huntington argued that the West's problem is not Islamic fundamentalism but Islam. It is beyond the scope of this entry to discuss the philosophical, historical, and empirical flaws of this argument. Suffice it to say that Islam, like any other major religion, is not a monolith and that it has given rise to different interpretations and practices at different times and in different places. Likewise, historically speaking, it has been compatible with different forms of government.

Looked at from a purely empirical point of view, what is claimed to be "Muslim exceptionalism" seems more like "Arab exceptionalism." Thus, Alfred Stepan convincingly argues that Islam is sometimes misleadingly equated with Arab culture. He observes that, while there are no democracies in the Islamic countries of the Arab world, about half of all the world's Muslims live in democracies, near-democracies, or intermittent democracies. He then analyzes the cases of Indonesia, Bangladesh, Pakistan, and Turkey as examples of such hybrid regimes. Vali Nasr describes the same cases (plus Malaysia) as examples of the "rise of Muslim democracy," which he compares with the early stages of Christian democratic parties in Europe. He states that, unlike hard-line Islamists, Muslim democrats have a pragmatic perspective and seek to build governing coalitions within a democratic arena whose bounds they respect.

Turkey, with its militantly secular constitutional and legal system and its more than 60 years of (albeit interrupted) democratic experience, is perhaps the leading example in this category. However, the extensive tutelary powers enjoyed by the military, and its less than desirable record on political rights (judged by the ease and frequency of party closures), allow us to classify it only as an electoral democracy. Thus, according to the *Freedom in the World 2009* report, it is an electoral democracy in the "partly free" category with a rating of 3 for both political rights and civil liberties, on a scale of 1 to 7, with 7 being the lowest score. According to the same survey, the major non-Arab, Muslim-majority countries have the following ratings (the first figure refers to the countries' score on political rights and the second on civil liberties): Albania (3 and 3), Bangladesh (4 and 4), Bosnia-Herzegovina (4 and 3), Indonesia (2 and 3), Malaysia (4 and 4), Pakistan (4 and 5), Senegal (3 and 3), and Turkey (3 and 3). Of these, Indonesia is placed in the "free" category and the rest in the "partly free" category, and all are rated, except for Malaysia and Pakistan, as "electoral democracies."

Although the Freedom House's threefold classification of "free," "partly free," and "not free" categories are not identical to Schedler's fourfold (liberal democracy, electoral democracy, electoral authoritarianism, and closed authoritarianism) or Diamond's sixfold (liberal democracy, electoral democracy, ambiguous regimes, competitive authoritarian, hegemonic electoral authoritarian, and politically closed authoritarian) typologies, there is, no doubt, a large area of overlap. Thus, liberal democracy clearly corresponds to the "free" category, while closed authoritarianism corresponds to "not free." On the other hand, Freedom House's partly free category includes both electoral democracies and electoral authoritarian regimes (or the four intermediate regime types in Diamond's still finer typology), and the borderline between the two may not always be clear. It may be said as a rule of thumb, however, that countries with a Freedom House score of 3 or 4 may be considered electoral democracies, while those with a score of 5 or 6 belong to the electoral authoritarian category.

Among the non-Arab, Muslim-majority countries, the former Soviet states occupy a special place in that they are all in the "not free" or "closed authoritarian" category, with the exception of Kyrgyzstan, which is rated partly free by the Freedom House (with scores of 5 and 4). Thus, Azerbaijan, Kazakhstan, and Tajikistan have scores of 6 and 5, while Turkmenistan and Uzbekistan are at the authoritarian end of the continuum with scores of 7 and 7. This can be explained by their strong totalitarian legacy. Iran (discussed above) and Afghanistan are the other two special cases.

Authoritarianism in the Arab World

The situation is not at all encouraging in the Arab core of the Middle East and North Africa. In 2000, among 16 countries, 10 (Algeria, Egypt, Iraq, Libya, Oman, Qatar, Saudi Arabia, Syria, Tunisia, and the United Arab Emirates)—or 12 (if we include two border states Mauritania and Sudan that have an Arab majority)—were in the "not free" category and only 6 (Bahrain, Jordan, Kuwait, Lebanon, Morocco, and Yemen) in the "partly free" category. During the decade, Bahrain, Jordan, and Yemen came into the "not free" category, and consequently, in 2009, Freedom House scores for the remaining partly free Arab countries were as follows: Kuwait (4 and 4), Lebanon (5 and 3), and Morocco (5 and 4). Even more interesting, Stepan observes that, while the 16 Arab countries are "underachievers" in electoral competitiveness (relative to what one would expect from their levels of gross domestic product per capita), the 31 Muslim-majority but non-Arab countries are "overachievers."

To explain "Arab exceptionalism" requires no less an effort than explaining the much more debatable "Muslim exceptionalism." It does not seem convincing to attribute the Arab democracy gap to certain inherent characteristics of the Arab political culture. Therefore, one should look for political, rather than ethnic or religious, particularities of the region to explain such exceptionalism. But exactly what kind of "political" particularities? Indeed, authoritarianism in the Arab world has given rise to an extensive literature. Although it is not possible to give the subject a full treatment here, the most common explanations can be summarized as follows.

One such explanation is related to the historical conditions under which Arab states emerged. As Stepan correctly states, many contemporary Arab states have relatively new and arbitrary boundaries cut out of the Ottoman Empire and were afterward occupied and colonized by European powers. Compounding this problem of state identity is the widespread appeal of pan-Arabism in many parts of the region, based on a common language and common religion. Against this background, it is remarkable that the new Arab territorial states have shown a high degree of resilience and durability. Egypt, with its strong sense of identity and tradition of statehood going back to several millennia, may be considered an exceptional case. Another related factor is the impact of the colonial or semicolonial experience in the political culture of Arab states, where independence from foreign rule was the dominant objective overshadowing concerns such as democracy and human rights.

A second explanation centers on the concept of "rentier state." Many Arab countries are rich in oil and natural gas reserves. The vast income deriving from these sources enabled the ruling elites to buy or co-opt some elements of the potential opposition as well as satisfy the basic material needs of a large majority of the population. This has had a retarding effect on the emergence of an independent bourgeoisie that is the driving force for democratization in Muslim as well as other countries. As Nasr says, Muslim democracy needs the bourgeoisie, and the bourgeoisie needs Muslim democracy. Nevertheless, the rentier state hypothesis cannot explain the prevalence of authoritarianism in oil-poor Arab states.

A third argument, perhaps the most convincing one, provides a cyclical explanation. Thus, the repressive policies of Arab authoritarian rulers against all opposition groups, secular and Islamist alike, and their refusal to even partially liberalize the system radicalize the opposition. The chief beneficiary of this process is the Islamist opposition, much stronger, more broadly based, and better organized than secular opposition groups. This, in turn, increases the perception of threat (real, exaggerated, or manipulated) on the part of the ruling groups and is used to justify even more repressive policies. This fear, although highly exaggerated and manipulated, is also vivid within the Turkish state elites (particularly the military and the judiciary), and it constitutes one of the most formidable roadblocks on the way to further democratization.

The feeling of an Islamist threat also leads to a paradoxical situation in that liberal or secular opposition groups that should normally have been on the side of democratic reforms also support repressive policies against the Islamist opposition. As Daniel Brumberg argues, the alliance between potential democrats and the authoritarian states in the Middle East is the reverse of the political reform process in Eastern Europe. In the latter cases, the not too democratic elites nonetheless found democratic procedures useful for dealing with the opponents of the regime, thereby leading to a "democracy without democrats." In the Middle East, by contrast, fear of Islamist victories has produced "autocracy with democrats," as potential democrats now actively support or at least tolerate autocrats as a lesser evil than an Islamist regime. A similar process is under way in Turkey in the ostensibly "social democratic" and militantly secularist RPP's (Cumhuriyet Halk Partisi) support for the military's harsh attitudes and manipulative maneuvers against the governing Justice and Development Party (Adalet ve Kalkınma Partisi; AKP), even though the latter can in no way be considered an Islamist or even a post-Islamist party.

The dilemma of the authoritarian rulers in the Middle East explains the zigzagged course of liberalization policies. It has been observed that periods of relative liberalization have often been followed by those of deliberalization. Although in some Arab countries, such as Jordan, Morocco, Algeria, Kuwait, and Egypt, controlled or semicompetitive elections have taken place, and moderate advances have been made in areas such as the freedom of the press, association, and assembly, none of these regimes has advanced beyond the level of electoral (or competitive) authoritarianism. Many studies of these cases are rather optimistic, probably reflecting the democratizing bias of their authors.

And yet it seems a fruitful line of inquiry to distinguish between the relatively more successful and less successful of these cases. Thus, Brumberg distinguishes between what he calls "dissonant" states and "harmonic" states. The dissonant states leave some room for competitive or *dissonant* politics. The examples are Morocco, Egypt, Jordan, Kuwait, and Lebanon. That such regimes, however autocratic, recognize and do not attempt to destroy societal pluralism creates an arena for competition and negotiation. The "harmonic" states, on the other hand, in their attempt to impose an often artificial unity through repression and cooptation, leave little room for negotiation and compromise and create a deadly game of "winner takes all." These are the states based on an ideology (be it Islam, Arab socialism, or Kemalism) and a strong sense of mission that attempt to mold the society according to their image of the "good society." John Waterbury characterizes such states as "ends-oriented" states, arguing that since the mission is considered sacred, debates over ends or means are seen as subversive or blasphemous.

The dissonant Arab states and those in the non-Arab parts of the Muslim world permitted a degree of electoral competition that, in turn, helped moderate the Islamist groups and integrate them into the political system. Turkey is the leading example of this process, where the earlier Islamist parties were transformed into a moderate, center-right, conservative democratic party (the AKP). Other examples include Bangladesh, Indonesia, Malaysia, and Pakistan. Similar processes are under way in Morocco, Jordan, Kuwait, and Lebanon. As Nasr observes, competitive elections push religious parties toward pragmatism and make other parties more sensitive to Islamic values in a game to win the middle. Electoral competition may eventually lead to a situation of "democracy without democrats," and this surely is much more preferable to that of "autocracy with democrats."

Conclusion

The early optimism and enthusiasm about the global resurgence of democracy have been replaced by a sense of greater realism. It is clear that authoritarian regimes, in one variety or other, have shown greater resilience and staying power than originally anticipated, using a skillful mixture of coercion, co-optation, divide-and-rule tactics, and selective democratic openings. However, just as the prospects for sustaining full-scale (or closed) authoritarian regimes in an age of democratization are not good, so are the prospects for electoral (or competitive) authoritarian regimes. The uneasy mixture of democratic and authoritarian elements creates an inherent source of instability, with the possibility of going in the direction of either electoral democracies or less competitive hegemonic authoritarian regimes.

Ergun Özbudun
Bilkent University
Ankara, Turkey

See also Democracy, Types of; Political Systems, Types; Totalitarian Regimes; Totalitarianism; Traditional Rule

Further Readings

Brumberg, D. (2002). The trap of liberalized autocracy. *Journal of Democracy, 13*(4), 56–68.
Carothers, T. (2002). The end of the transition paradigm. *Journal of Democracy, 13*(1), 5–21.
Collier, D. (Ed.). (1979). *The new authoritarianism in Latin America*. Princeton, NJ: Princeton University Press.
Diamond, L. (2002). Elections without democracy: Thinking about hybrid regimes. *Journal of Democracy, 13*(2), 21–35.
Hermet, G., Rose, R., & Rouquié, A. (Eds.). (1978). *Elections without choice*. New York: Wiley.
Levitsky, S., & Way, L. A. (2002). Elections without democracy: The rise of competitive authoritarianism. *Journal of Democracy, 13*(2), 51–65.
Linz, J. J. (1964). An authoritarian regime: Spain. In E. Allardt & Y. Littunen (Eds.), *Cleavages, ideologies, and party systems: Contributions to comparative political sociology* (pp. 291–341). Helsinki, Finland: Westermarck Society.
Linz, J. J. (2000). *Totalitarian and authoritarian regimes* (with a major new introduction). Boulder, CO: Lynne Rienner.
Nasr, V. (2005). The rise of Muslim democracy. *Journal of Democracy, 16*(2), 13–27.
O'Donnell, G. (1973). *Modernization and bureaucratic authoritarianism: Studies in South American politics*. Berkeley: Institute of International Studies, University of California.
Posusney, M. P., & Angrist, M. P. (Eds.). (2005). *Authoritarianism in the Middle East: Regimes and resistance*. Boulder, CO: Lynne Rienner.
Salamé, G. (Ed.). (1994). *Democracy without democrats: The renewal of politics in the Muslim world*. London: I. B. Tauris.
Schedler, A. (2002). Elections without democracy: The menu of manipulation. *Journal of Democracy, 13*(2), 36–50.

Schedler, A. (Ed.). (2006). *Electoral authoritarianism: The dynamics of unfree competition.* Boulder, CO: Lynne Rienner.
Stepan, A. (with Robertson, G. B.). (2003). An "Arab" more than a "Muslim" democracy gap. *Journal of Democracy, 14*(3), 30–44.

AUTHORITY

See Legitimacy

AUTONOMY, ADMINISTRATIVE

Within democratic systems, public organizations carry out a large number of functions, ranging from provision of education, health, and national security and regulation of financial markets to ensuring equal treatment for all. These bodies perform these functions because delegation is imperative: Elected politicians lack the time, expertise, and resources both to enact and implement laws and to delegate these tasks to public organizations. To perform their tasks, these bodies need a certain level of autonomy from democratic oversight. The laws they implement often contain goals, such as "national security" and "social welfare," that need further specification before they can be attained. Also, several different approaches and policy instruments to complex social problems may exist, and the choice of approach or policy instrument requires experience and knowledge of these social problems. Public organizations are staffed by civil servants with such experience and knowledge, but they need some level of autonomy to analyze social problems and to choose the best possible course of action for addressing these. Administrative autonomy revolves around a delicate balance. On the one hand, if public organizations lack administrative autonomy, they run the danger of being micromanaged by politicians. On the other hand, if there is too much administrative autonomy, politicians and voters run the risk of having created public organizations that are nonresponsive to democratic preferences.

Administrative autonomy is therefore a core concept in the study of public administration. At its heart lie democratic-theoretical questions concerning the relationship between voters, politicians, and civil servants as well as rational-instrumental issues concerning the management and design of effective and efficient public policies. As a theoretical concept, administrative autonomy represents an abstract social construct. Under this lemma, the background of the concept of administrative autonomy within the field of public administration is discussed first. Next, the various attributes and dimensions of the concept of administrative autonomy are described. Finally, an overview of theories of administrative autonomy is provided.

The Concept of Autonomy in the Study of Public Administration

In the field of public administration, the concept of administrative autonomy has been enjoying the close attention of many scholars for more than several decades (see, e.g., Daniel Carpenter, 2001; Peter Clark & James Wilson, 1961; Anthony Downs, 1967; Philip Selznick, 1957). Administrative autonomy is considered a sine qua non for nothing less than the "survival" of a public organization. To survive competition with other public organizations over scarce resources, the attaining of a substantial degree of administrative autonomy is seen as a critical goal: "Autonomy gives an organization a reasonably stable claim to resources and thus places it in a more favorable position from which to compete for those resources" (Clark & Wilson, 1961, p. 158). Administrative autonomy as the lifeblood of successful public organizations is still a vivid argument in contemporary public administration scholarship. In his study of the modernization of the American bureaucracy during the late 19th and early 20th centuries, Carpenter (2001) ascribes the success of various agencies, among them the U.S. Postal Service and the U.S. Food and Drug Administration, to their leaderships' capabilities to forge autonomy for their organization.

During the 1990s, we witnessed a renewed interest in the concept of administrative autonomy. Three developments were important drivers of this interest. First, starting in the late 1980s and the beginning of the 1990s, by setting up "quasi-nongovernmental organizations," that is, quangos, governments worldwide embarked on the path of restyling their central governmental apparatus according to one of the core prescriptions of the new public management (NPM) paradigm. In traditional Western parliamentary systems, governments were confronted with trade-offs between giving autonomy to parts of their integrated central administrative apparatuses and to retain control over the newly created autonomous bodies.

The second driver of the renewed attention for the concept of autonomy is the wide diffusion and creation of independent regulatory authorities (IRAs). The intellectual foundations for the creation of IRAs stem from the same source as that of NPM described above. The financial and economic reforms of the 1980s and 1990s resulted in the liberalization of markets and the breaking down of public monopolies. Deregulation led to reregulation as contracts and property rights within these newly created markets needed public supervision. Legislatures entrusted the protection of property rights and the enforcement of contracts on these new markets to independent regulatory agencies. Their autonomy from government is considered a necessary condition for guaranteeing the credible commitment of the government to not renege on the liberalization of these markets. Consequently, the creation of IRAs brought immediately to the fore the question of who controlled these highly autonomous bodies.

The third and final driver is the emergence of neo-economic theories within public administration and political science. Neo-economic theories entered the fields of political science and public administration through the latter's adaptations from economic and legal approaches to contract theories. Neo-economic theories first made school within American political science. The warm reception of the principal–agency framework is partly the result of the decades-old American interest in Congress's capability to control the bureaucracy. Principal–agency and transaction costs theories introduced a new conceptual tool kit into public administration and concepts such as "contract," "information asymmetry," "monitoring," and "costs" started to dominate our talk and thinking about politics and administration. Neo-economic theories soon crossed the Atlantic and assumed a prominent place within European studies on regulation, Europeanization, and quangocratization. This resulted in the growing attention on the concept of administrative autonomy.

The study of administrative autonomy is colored by sociological and neo-economy–inspired political science theories. Whereas sociological theories treat administrative autonomy as an important asset of public organizations to successfully function within their own environment, economic-political science theories consider the autonomy of public organizations as a potential liability. We will see that these perspectives constitute the basic view within two of the three main sets of theories of administrative autonomy.

The Anatomy of Administrative Autonomy

Administrative autonomy is an abstract concept that is not directly observable. In this section, we will present two dimensions of administrative autonomy that are commonly used in studies of this concept.

De Facto and De Jure Autonomy

The first distinction is between formal or de jure autonomy and actual or de facto autonomy of public organizations. De jure autonomy refers to the formal-legal prescriptions concerning the relationship between politicians and public organizations. The laws that created the body (or the laws and regulations that delegate specific tasks to public organizations) state how a political principal and its administrative agent should interact in *theory*. As such, these laws spell out the rules of the game that should be obeyed when these actors interact with each other. To a certain extent, the de facto or actual autonomy of a public organization is reflected by its formal autonomy. A public organization that is not part of a ministerial department, for example, will have more formal administrative autonomy than a unit of a ministerial department.

However, the correlation between de jure and de facto administrative autonomy is not always strong. George Krause and James Douglas (2005), for example, found that it was not the formal administrative autonomy of three different U.S. economic agencies that affected the outputs of their economic forecasts but the reputation of the economic professionals who worked within these agencies. Martino Maggetti (2007) found that a higher level of formal administrative autonomy is not a sufficient nor necessary condition for higher levels of de facto autonomy. These findings point at the existence of a discrepancy between rules and competences on paper and rules and competences in practice. In the words of Kutsal Yesilkagit (2004),

> Rules, procedures, and competences may seem clear on paper as they inform all actors about the formal intentions of the designers. However, [after the body is created] public organizations leadership can interpret the rules in a different way than politicians had in mind when they designed the rules. (p. 535)

Dimensions of Autonomy

Koen Verhoest, Guy B. Peters, Geer Bouckaert, and Bram Vermeulen (2004) distinguish between seven organizational domains and related dimensions of autonomy. The first dimension is *managerial* autonomy and refers to the extent in which managers may decide on the allocation of the input (e.g., budget) and resources (e.g., personnel) of public organizations without prior approval from a supervisory body (i.e., a minister). Second, *policy* autonomy designates the discretionary space within which managers of public organizations can decide on the quantities of the goods their organizations produce, choose from among the policy instruments they have to their avail, and arrange the internal procedures within their organizations. Managerial and policy autonomy hence correspond to the discretionary authority of public managers: a space within which managers enjoy some leeway to manage their organizations as they see fit.

External stakeholder may actively limit the discretionary space of the management of public organizations. Administrative autonomy is then a function of the exemption from constraints on the actual use of decision-making competencies that external stakeholders—that is, the ministers—grant these organizations. In democratic systems, central governments always retain certain powers to fine-tune the amount of decision-making competences of public organizations. This fine-tuning may take place through the adjustment of an organization's structural autonomy, legal autonomy, interventional autonomy, and financial autonomy. *Structural* autonomy is the extent to which the head of public organizations is exempted from direct supervision of elected officials—that is, the minister. If the head of a public organization reports to a governing board and not directly to the minister, then the structural autonomy of public organizations may be considered as high. The *financial* autonomy of a public organization is high when it is entitled to draw its own financial resources from the revenues of the services the organization delivers; it is low when the revenues of the organization come from the departmental budget. With regard to the *legal* dimension of autonomy, Verhoest et al. take the legal basis of the public organization as the measuring rod. If a public organization lacks legal personality or it is established by a statutory instrument instead of a statute, the position of the public organization vis-à-vis the legislature or the cabinet will be weak, as the latter can change the legal status of the public organization with lesser (legislative) costs than when public organizations was founded on primary law or a statute. Finally, *interventional* autonomy is the extent to which a public organization is exempt from *ex post* reporting, evaluation, and audit requirements. Table 1 gives a brief schematic overview of the dimensions of autonomy and the substances to which autonomy and/or exemptions refer.

Theories of Administrative Autonomy

Until now we have discussed the background and various attributes of the concept of autonomy. Here we will focus on the factors that affect levels of administrative autonomy. In research as well as in practice, we are primarily interested in why some public organizations have more administrative autonomy than others and what these organizations (can) do with their autonomy in public life. Overlooking academic research on administrative autonomy over the past decades suggests that there exist three sets of theories of administrative autonomy: (1) *external-control* theories, (2) *internal-administrative* theories, and (3) *dynamic-adaptive* theories.

Table 1 Dimensions of Autonomy

Type of Autonomy	*Autonomy Dimensions*	*Substance of Autonomy*
Level of decision-making competencies (discretionary powers)	Managerial autonomy	Decisions on inputs and allocation of resources
	Policy autonomy	Decisions on policy instruments, production of goods
Exemptions from constraints on decision-making competencies	Structural autonomy	Exemptions from direct hierarchical supervision
	Legal autonomy	Legal basis of public organization
	Financial autonomy	Source of agency funding
	Interventional autonomy	Exemption from *ex post* controls

Sources: Adapted from Verhoest, K., Peters, G. B., Bouckaert, G., & Vermeulen, B. (2004). The study of organisational autonomy: A conceptual overview. *Public Administration and Development, 24*(2), 101–118; Christensen, J. G. (2001). Bureaucratic autonomy as a political asset. In B. G. Peters & J. Pierre (Eds.), *Politicians, bureaucrats and administrative reform* (pp. 119–131). London: Routledge.

External-Control Theories of Administrative Autonomy

External-control theories of administrative autonomy claim that the administrative autonomy of a public organization is a function of the level of pressures exerted on the organization by political stakeholders in the environment of the public organization. The theories that are grouped under this heading have in common that they apply a (version of) principal–agency theory of political control. They assert that the degree of administrative autonomy of public organization depends on the instruments and resources that political principals, such as Congress or the president (in U.S. studies) and parliaments or cabinets and ministers (in studies on parliamentary systems), have to control the actions and outputs of public organizations. These instruments are usually divided between *ex ante* controls, such as agency design, administrative procedures, and *ex post* controls, such as audits, reporting requirements, and (threats of) budgetary cuts.

External control is not limited to political principals. Next to formal political actors, interest groups too may act as principals of public organizations. Interest groups are key players in the politics of delegation and administrative design, as they, perhaps more than elected politicians, care about the control regime of public organizations as these organizations are essential to the policy process and the distribution of public services to their members. Political control is therefore not an

> issue . . . of congressional dominance or presidential control, but rather of shared authority and the ability to influence policy in one direction or another, a process whereby interest groups' ability to access and shape policy both directly and indirectly is evident and must not be overlooked. (Scott Furlong, 1998, p. 61)

In other words, political control is a form of "joint custody" of president, congress, interest groups, and the courts, and administrative autonomy is the resultant of the strategies pursued by multiple interacting actors.

Internal-Administrative Theories of Administrative Autonomy

In contrast to external control theories, internal-administrative theories of administrative autonomy put public organizations and their interests, capacities, identity, and mission central to their explanatory models. The main thesis of this line of research is that the administrative autonomy of public organizations is a function of the organization's capabilities to forge its own autonomy from its direct environment. Whereas we could say that in the former perspective, the organization is considered as a passive taker of the discretion that its environment grants to itself, in this perspective, public organizations are considered as active seekers or forgers of autonomy (Carpenter, 2001).

As suggested by Wilson (1989, pp. 188–192), two sets of internal-administrative factors are deemed crucial for achieving autonomy. One is "political legitimacy" (Carpenter, 2001)—that is, the belief and reputation that agencies can deliver the services that are "worthwhile to some group with influence over sufficient resources to keep it alive" (Downs, 1967, p. 7). Agencies should "impress those politicians who control the budget that its function generate political support" (p. 7) from which politicians can benefit in terms of their reelection or strengthening of their party political base. Second, administrative autonomy requires the "organizational capacities" to deliver the benefits to their supporters: It must have the staff with the right skills and dedication to fulfill the organization's ultimate goal of securing a stable and strong autonomous position within the public space. In line with this, Selznick (1957) saw carefully designed recruitment instruments as a prerequisite for autonomy. The leadership of an organization, if it wanted to pass the initial survival threshold and maneuver through critical periods ahead, had to carefully select, first, the "social base," that is, those segments of "the [organization's] environment to which operations will be oriented" (Selznick, 1957, p. 104).

Dynamic-Adaptive Theories of Administrative Autonomy

The third set of theories of administrative autonomy commonly claims that the administrative autonomy of public organizations is a function of a complex pattern of interactions between public organizations and its (political) environment. The interactions consist of a mix of stimuli or signals and responses that traverse from various venues in the environment to a public organization as well as from agencies to the political actors in the environment. "[C]ausality" between control and autonomy "does not just flow from each democratic institution towards the administrative organization as assumed in many conceptualizations of political control theories, but . . . also flows from agency to political principals" (Krause, 1996, p. 1089).

According to Dan Wood and Richard Waterman, stimuli are discrete events (e.g., a crisis, the election of a new president), event progressions (e.g., yearly budgetary appropriations, congressional oversight hearings), and tonal stimuli (i.e., changes in the tone of political–bureaucratic relationships), and the responses are a function of the agencies' technological, rational, and political assets. Unanticipated behavior, a function of the general uncertainty in any policy environment, also influences autonomy. Policies are made in ambiguity; agencies do not only often poorly anticipate possible events, but they are also often not capable of making sense of the stimuli that reach them, causing public organizations to display unanticipated response behavior.

Conclusion

Administrative autonomy is a core concept in the study of public organizations. As the concept stands, it is studied in close relation to democratic-theoretical notions such as legitimacy, public accountability, and democratic control. In the preceding sections, we have discussed the various attributes of this concept and theories of administrative autonomy. We have seen that research on this topic is substantive. However, there are still aspects of administrative autonomy that have not yet been fully studied. Here, three lines of research that are worthy to be explored by future research are discussed.

In the studies discussed above, administrative autonomy is treated as the dependent variable. However, from a democratic perspective, the degree of administrative autonomy of public organizations is particularly important for the extent it enables or restrains public organizations in their operation. Public organizations that are more or less autonomous provide public services and regulate affairs across a variety of societal domains and may deeply affect the lives of citizens. So one line of research that deserves to be further elaborated is on the relationship between administrative autonomy and performance or policy impact. The study of administrative autonomy may be connected to existing studies on bureaucratic responsiveness. These studies examine the impact interest groups have on decision-making processes of public organizations but leave the relationship with administrative autonomy somewhat unexplored.

Second, an underexamined aspect of administrative autonomy is of a more normative nature, namely, the discussion from which stakeholders and public organizations *should* function autonomously. Public organizations are created to implement democratically chosen public policies. So, autonomy from government will never be absolute, as public organizations have been entrusted with the implementation of collectively agreed-on public policies. Furthermore, public organizations also interact with and are responsive to a number of nongovernmental organizations. So, public organizations will not be entirely autonomous from nongovernmental stakeholders as well. The focus within existing research has predominantly been on the varying degrees of autonomy from governmental and nongovernmental stakeholders, without questions regarding the desirability of varying degrees of autonomy from different stakeholders.

Finally, the globalization of the economy and technological developments increasingly blur the lines between the international order and domestic administration. We move toward a postnational bureaucracy that is hallmarked by the emergence of dense international networks of civil servants. Administrative autonomy becomes situated into a new context as the environment of public organizations becomes more and more transnational. The consequences of globalization of the executive branch force on us new ways of thinking about the relationship between politics and administration. Here, too, administrative autonomy will be a core concept.

Kutsal Yesilkagit
Utrecht University
Utrecht, Netherlands

See also Administration; Administration Theory; Autonomy, Subnational; Bureaucracy; Sovereignty

Further Readings

Carpenter, D. P. (2001). *The forging of bureaucratic autonomy: Reputations, networks, and policy innovation in executive agencies, 1862–1928*. Princeton, NJ: Princeton University Press.

Christensen, J. G. (2001). Bureaucratic autonomy as a political asset. In B. G. Peters & J. Pierre (Eds.), *Politicians, bureaucrats and administrative reform* (pp. 119–131). London: Routledge.

Clark, P. B., & Wilson, J. Q. (1961). Incentive systems: A theory of organizations. *Administrative Science Quarterly, 6*(2), 129–166.

Downs, A. (1967). *Inside bureaucracy*. Boston: Little, Brown.

Furlong, S. R. (1998). Political influence on the bureaucracy: The bureaucracy speaks. *Journal of Public Administration, Research and Theory, 8*(1), 39–65.

Krause, G. A. (1996). The institutional dynamics of policy administration: Bureaucratic influence over securities regulations. *American Journal of Political Science, 40*(4), 1083–1121.

Krause, G. A., & Douglas, J. W. (2005). Institutional design versus reputational effects on bureaucratic performance: Evidence from US government macroeconomic and fiscal projections. *Journal of Public Administration Research and Theory, 15*, 281–306.

Maggetti, M. (2007). De facto independence after delegation: A fuzzy-set analysis. *Regulation & Governance, 1*(4), 271–294.

Maor, M. (2007). A scientific standard and an agency's legal independence: Which of these reputation protection mechanisms is less susceptible to political moves? *Public Administration, 85*(4), 961–978.

Selznick, P. (1957). *Leadership in administration: A sociological interpretation*. New York: Row, Peterson.

Verhoest, K., Peters, G. B., Bouckaert, G., & Vermeulen, B. (2004). The study of organisational autonomy: A conceptual overview. *Public Administration and Development, 24*(2), 101–118.

Wilson, J. Q. (1989). *Bureaucracy: What government agencies do and why they do it*. New York: Basic Books.

Wood, D. B., & Waterman, R. (1991). The dynamics of political control of the bureaucracy. *American Political Science Review, 85*, 801–828.

Yesilkagit, K. (2004). Bureaucratic autonomy, organizational culture, and habituation: politicians and independent administrative bodies in the Netherlands. *Administration & Society, 36*(5), 528–552.

AUTONOMY, SUBNATIONAL

The concept of autonomy is polysemous in scientific literature and can be used in very different contexts. When used with the word *subnational*, the concept refers to the degree of autonomy in governments below the level of central government, with the exception of local governments. Examples include the Canadian provinces, the U.S. states, the Spanish autonomous communities, the regions and communities in Belgium, the German and Austrian *Länder,* and the French departments and regions. Autonomy refers to the degree of decisional autonomy, sovereignty or self-government within a nation-state, enjoyed by these subnational governments.

This entry first discusses subnational autonomy and the structure of the state. Second, it describes the power, resources, and influence of subnational governments. Regionalization is discussed in the third section, the fourth section deals with multilevel governance and subnational autonomy, and the final one examines minority nationalism and subnational autonomy

Subnational Autonomy and the Structure of the State

Because nation-states vary in their structure, subnational governments also differ from country to country. In many cases, vast differences are found within a single country. This diversity results in a strong asymmetry between subnational governments in terms of their autonomy. The autonomy of subnational governments is based on several factors, including the institutional setting of a nation-state and the constitutional powers devoted to a subnational government. The more a country is constitutionally decentralized, the greater is the extent to which subnational governments have legislative powers and thus autonomy.

In unitary states such as Denmark or Israel, subnational governments have little or no autonomy. A unitary state is a state governed as one single unit in which the central government is the decision center. Subnational and local governments exercise only the power that the central governments choose to delegate.

In the case of a devolved or decentralized state, subnational governments have more autonomy. A decentralized state is generally a former unitary state such as France. The authority and responsibility for some public functions have been transferred from the central government to the regional government. A devolved state is a centralized state such as the United Kingdom (UK), where subnational governments have a degree of autonomous power devolved from the central government. A devolved government cannot challenge the constitutionality of central government's law. The power given by the central government can be revoked or reduced. For example, the Northern Ireland Assembly has been suspended many times by London since its creation. In theory, decentralized and devolved states are different, but in practice, the differences are very thin.

A federal state is a type of sovereign state in which sovereignty is constitutionally divided between a central government and the subnational governments (e.g., the Canadian provinces or the German *Länder*). To qualify as a federation, a minimum requirement is that the constitutional powers of the subnational governments cannot be changed unilaterally by the central government. In federal systems such as Germany, Canada, the United States, or Belgium, subnational governments have considerable autonomy. In Canada, for example, provinces are constitutionally responsible for health care, education, culture, and municipalities. They can also act on economic development, justice, environment, and so on. In Belgium, the autonomy of communities and regions goes even further. Since 1993, Belgium has been, according to the first article of its Constitution, "a federal state composed of communities and regions." The constitutional revision of 1993 permits the regions and communities to become real international actors. This includes the power of representation and power to sign treaties with sovereign states. Since the revision of the Constitution in 1993, the organization of Belgium's international relations is fundamentally adapted to the federal state structure. The autonomy of the Belgian substate actors, with regard to external policy, is unique in the world. Its exceptional nature arises from the recognized constitutional principle of *in foro interno, in foro externo*. On top of that, there is an *absence of hierarchy* between different levels of administration (see Table 1).

Subnational Governments Power, Resources, and Influence

The autonomy of subnational governments is also related to the resources that these actors can mobilize. Some subnational governments such as those

Table 1 Typology of Democracies and Cases in Relation to the Autonomy of Subnational Governments

Unitary States	*Decentralized and Devolved States*	*Federal States*
Denmark	France	Belgium
Finland	Netherlands	Canada
Greece	Portugal	Germany
Ireland	Great Britain	United States
Sweden	Spain	Austria
Israel	Italy	India

in Quebec, Catalonia, Flanders, California, or Bavaria possess more resources than many sovereign states. Let us compare subnational governments with the 192 members of the United Nations. In terms of population, 40 states have a population of less that 1 million and around 13 have between 1 million and 2 million. In the United States alone, 43 states have a population greater than 1 million, and 36 have a population of more than 2 million, according to U.S. Census Bureau data for 2008.

In 2008, according to the World Bank, nominal 2008 gross domestic product (GDP) estimates for 191 nation-states and territories, California would rank among the top 10 in the world. The top 11 American states would rank among the first 25 in the world, the top 26 in the United States are among the top 50 globally, and all 50 states are among the top 86 nation-states. The state of California could even be a member of the G8 (Group of 8; or the G20 [Group of 20]) since its GDP is greater than those of Canada and Spain. The annual budget of California exceeds that of Mexico and numerous sovereign states. Thus, subnational governments are among the most important economic actors in the world.

Regionalization

Since World War II, authority has shifted away from the central state to international organizations, such as the European Union or the World Trade Organization, and to regions and municipal governments. Regionalization represents a transformation of regional authority where the nation-state creates or reinforces the mesogovernment level between the central state and the local level. If Canada and the United States have been federations for years, in Europe regionalization has gone hand in hand with European integration.

After World War II, only Germany, Switzerland, and Austria were considered decentralized federal states. Since the 1950s, of the 27 countries that are now part of the European Union, including Belgium, Spain, Italy, the UK, and France, 19 have seen the rise of a stronger and deeper regional or subnational government. These governments now handle a wide variety of policy responsibilities. Since 1950, new regional governments have been created in 14 countries of the European Union (EU). Eighteen regions in 7 countries, excluding the Belgian regions and communities and the Spanish autonomous communities, have also been given a special autonomous or different status. In 1950, only 5 countries had directly elected regional assemblies covering the countries as a whole; by 2010, 16 countries had such assemblies. Only 2 countries in Europe with a population of more than 2.5 million, Sweden and Bulgaria, have not followed this trend. In Sweden, decentralization reinforced the role of municipal governments instead.

These meso or regional governments are a fairly new thing in many countries with diverse political history and constitutional setting. These new regional governments are reshaping politics, the distribution of power between the central government and the subnational units, and accountability.

According to some researchers, the driving force behind the rise of these regional governments is related, in part, to globalization. According to some scholars, the nation-state today is too small for the big problems of the world and too big for the small ones. In this context, regional governments are better placed to deliver important public goods.

Other scholars identify the rise of regions with the decline of the authority of the nation-states. In Europe, European integration has reinforced regionalization and decentralization. European integration has noticeably weakened the centralized uniform states with the introduction of the principle of subsidiarity in the Treaty of Maastricht of 1992.

There are also more technical reasons for regionalization or decentralization. For some scholars, the virtues of centralization and decentralization differ from one policy area to the other. The advantage of centralization is obvious: economies of scales and, in theory, greater equality between the citizens of one country. The virtues of decentralization are different. Some scholars suggest that regionalization and decentralization increase the efficiency of public policy. The fundamental principle of public goods analysis is that a jurisdiction should encompass those who are positively or negatively affected by a specific policy. Because the public goods differ, the scale at which they are most efficient also differs. Some policies, such as urbanism or fire protection, are best handled at the local level. Some policies, such as education or hospitals, are best handled at the regional level, and others such as national transportation, infrastructure, trade, and national security are better handled at the national level.

Some believe that political participation is better and more effective if the decision center is closer to voters. Decentralized governments are closer to the citizens and therefore have better knowledge about what they want and need. Regionalization thus contributes to better democracy because it multiplies opportunities for citizens to influence governments.

According to some researchers, regionalization and decentralization are drivers for better democracy. Elected officials in democratic states may shift decision making away from the central state if it can lead to more efficient decision making and can attract more votes. By contrast, authoritarian states tend to centralize the decision-making process. At best, authoritarian regimes will deconcentrate powers in local outposts so that they can provide more reliable information about local politics and implement more effective central policy. This fact explains why the democratization of southern, central, and eastern Europe went hand in hand with decentralization and subnational autonomy. The so-called third wave of democracy generated more decentralization. Regionalization and decentralization could thus, in theory, lead to more effective public administration and better democracy.

Finally, some specialists think that decentralization comes primarily from the public finance crisis. Regional governments are, in theory, more in phase with the real needs of the population and thus allocate resources more effectively. With the crisis of the public debt in numerous countries, decentralization is a way for the central state to transfer some responsibility to regions in order to reduce the pressure on public finance at the central level.

Multilevel Governance

With regionalization and decentralization, the responsibilities of substate governments have expanded considerably. Subnational governments have more power concerning economic development, education, health care, environment, public transport, and so on. Each of these public policy domains requires coordination among governments at diverse levels. Because no government is an island, increased globalization and internationalization have reinforced the need to coordinate local, regional, national, and international policies as the responsibilities of subnational governments expand.

The rise of regional government and subnational autonomy has a big impact on multilevel governance. Decisions taken at one level of government affect the decisions of the other level of government. Most policy thus require some form of coordination among international, European, national, regional, and local governments. We know that, for example, according to a United Nations estimate, between 50% and 70% of all adaptation and mitigation measures against climate change will call for implementation by the regions or subnational governments

The concept of multilevel governance was created within the framework of the EU to explain the relation between the various levels of government in EU policy making. Multilevel governance means that there are multiple actors from various level of government interacting to negotiate and implement public policy coming from the EU. The multilevel governance approach illuminates the interdependence between the local, regional, national, and international levels of authority.

At first, multilevel governance was developed for the study of the EU. Now, however, it is applied in various situations because virtually all government activities today are affected by the competence of at least one intergovernmental organization, and frequently many more. In this way, in the context of international organizations and international conferences, themes are dealt with that relate to education, public health, cultural diversity, the environment, business subsidies, the treatment accorded to investors, the removal of nontariff barriers, barriers to agriculture, to services, and so on. This phenomenon is magnified in Europe by the process of European integration and in North America by the North American Free Trade Agreement. Likewise, enlarging the scope of international issues means that all government departments have activities that are internationalized. This situation makes it harder for a country's ministry of foreign affairs to centralize the decision-making process.

In this context, subnational governments have become more aware that their political autonomy and their sovereignty—or, in other words, their ability to formulate and implement policy—are subject to negotiation in multilateral fora. Thus, since the 1960s, there has been a noticeable increase in the number of subnational governments that are interested and participate actively in international affairs.

In the United States, for instance, only 4 states had representative offices in other countries in 1970 versus 41 states with 245 representative offices in 34 countries in 2008. In Germany, the *Länder* have set up some 130 representative offices since 1970, of which 21 are located in the United States. In 2009, Quebec, one of the pioneers in the field, had some 28 representative offices around the world. In Spain, the autonomous region of Catalonia operates some 50 representative offices abroad, and in 2010 the Flemish government has more than 100 political, economic, and economic representations abroad. This phenomenon is also evident in Japan and many other countries.

When considered from the perspective of theories of federalism and international affairs, the two conceptions conflict: the centralizing school and the school devoted to multilevel governance. From the perspective of the centralizing approach, some scholars believe that a monopoly of foreign affairs is a "minimum power" of all central governments. In his landmark study, Kenneth Clinton Wheare highlighted the negative consequences of unbundling central control over foreign affairs for the national interest and for the functioning of the international system. To some scholars, international relations are at the heart of federal regimes. Centralization of the foreign affairs power is a requirement of international law because a centralized political system is a necessary condition for states to be able to play the role they are assigned in international law and practice. In essence, without the existence of a central government that has a plenary authority on its territory in relation to foreign affairs and the ability to participate in international relations and to enforce international obligations in the domestic order, interstate relations can only be seriously compromised. If the power of codecision is granted regarding treaty making, there would be a risk of paralyzing a state's foreign affairs. Every player would have a veto, resulting in harm to the state's image in the international arena.

Supporters of the concept of multilevel governance take a different view. According to some scholars, diplomacy or foreign policy cannot be considered a monopoly of the central government. Substate government will always have an important role, even if it is only for the purposes of implementing international agreements concluded by the central government. In addition, giving a monopoly over foreign affairs to central governments in federal regimes puts at risk the distribution of powers between the different orders of government for the benefit of central authorities. There are many examples of federal regime that must operate with important constitutional limitations on their powers in foreign affairs.

Foreign policy should be thought of as a complex system where different actors within the federal regime structure work with each other. Those who favor a multilevel governance approach thus maintain that "obligations of cooperation" exist between central governments and substate actors. To implement a coherent foreign policy, it is important to consult substate actors and, indeed, give them an important role by means of intergovernmental mechanisms, so that they can participate actively in the country's treaty-making process. According to this view, regional integration, multilateralism, and globalization have thus rendered centralist theses obsolete. The requirements of cooperation between the different orders of government are more and more important, and it is for this reason that one notices a considerable increase in executive federalism or intergovernmental relations in respect of the conclusion of international treaties in federal regimes.

Subnational Autonomy and Minority Nationalism

In multinational states, such as Canada, Spain, or Belgium, decentralization and subnational autonomy are also seen by many scholars as a way to reduce the likelihood of secession or political conflicts. Decentralization is a way to accommodate substate nationalist movements such as those in Catalonia, Flanders, or Quebec. The federalization or decentralization of multinational states is seen as a way to prevent the breakup of a country while permitting political and cultural autonomy of minority nations.

Many scholars have argued that some measure of autonomy or form of self-government will satisfy the majority of people within a minority nation. In practice, that would mean that a minority nation would, to some extent, govern itself while also participating in the nation-state institutions. Federalization of nation-states does not guarantee the end of the national question in multinational countries. The recent experience of Belgium with the Flemish movement, that of Scotland, or even those of Catalonia and Quebec confirm that political conflicts will remain, but it is hard to see any form of successful accommodation in multinational countries that does not include some elements of federalism and subnational autonomy.

Some researchers have reservations about granting subnational autonomy to accommodate minority nationalism. One problem is that federalism tends to accentuate the differences between the majority and the minority nation within a nation-state. The creation of an autonomous subnational government at the regional level gives the minority nation new tools to formulate new demands for more autonomy, which is seen as a never-ending story. Granting some form of self-government for minority nations tends to heighten and politicize the minority nation's self-consciousness and to facilitate the rise of a self-conscious intelligentsia. In the end, federalization is seen as a last-resort solution that risks fueling, rather than appeasing, minority nationalism. The territorialization of minority nations also risks reducing the cross-cutting interactions between the majority nation and the minority one. Giving some form of self-government to a minority nation holds the potential to polarize politics at the national level. It is thus very difficult or even impossible to create a national consensus on national issues.

Stéphane Paquin
Université de Sherbrooke
Sherbrooke, Quebec, Canada

See also Decentralization; European Integration; Federalism; Multilevel Analysis

Further Readings

Aldecoa, F., & Keating, M. (1999). *Paradiplomacy in action: The foreign relations of subnational governments*. London: Frank Cass.
Balme, R. (1996). *Les politiques du néo-régionalisme: Action collective régionale et globalization* [The policies of neo-regionalism: Regional collective action and globalization]. Paris: Economica.
Fry, E. H. (1998). *The expanding role of state and local governments in U.S. foreign affairs*. New York: Council on Foreign Relations Press.
Greer, S. L. (2007). *Nationalism and self-government*. Albany: New York University Press.
Hooghe, L., & Marks, G. (2001). *Multi-level governance and European integration*. Lanham, MD: Rowman & Littlefield.
Jeffrey, C. (Ed.). (1997). *The regional dimension of European Union: Towards a third level in Europe?* London: Frank Cass.
Keating, M. (1996). *Nations against the states: The new politics of nationalism in Québec, Catalonia and Scotland*. New York: St. Martin's Press.
Lynch, P. (1996). *Minority nationalism and European integration*. Cardiff, UK: University of Wales Press.
Moreno, L. (2001). *The federalization of Spain*. London: Frank Cass.
Paquin, S. (2004). *Paradiplomatie et relations internationals: Théorie des stratégies internationales des regions face à la mondialisation* [Paradiplomatic and international relations: Theory of international strategies of regions in globalization]. Brussels, Belgium: Peter Lang.

AUTONOMY AND SOVEREIGNTY

See Sovereignty

B

Balance of Power

The balance of power is, arguably, the most central theoretical concept in the study of international relations, yet it is also one of the most controversial concepts in the field. It is a significant concept because it underpins how theorists of a realist predisposition think about international relations, and realism is one of the most widely acknowledged general theories in the discipline. The concept is also controversial because realism, in general, and the balance of power, in particular, have always been fundamentally challenged by other schools of thought. Advocates of the balance of power argue that it helps account for the most fundamental features of international relations and, in particular, the survival and stability of any system of independent states. Critics, by contrast, attack the concept by insisting that it is either incoherent, thereby promoting a spurious understanding of international relations, or dangerous, because it promotes policies that can lead to international tensions and often war. This entry first examines the balance of power metaphor and describes its historical development. It then discusses the roles the concept has in contemporary thought, especially in the work of Hans J. Morgenthau and Kenneth N. Waltz.

The controversy that surrounds the balance of power is not a recent phenomenon. It can be traced back for several centuries, although there is no authoritative genealogy of the concept. Debates about the balance of power became a regular feature of the theory and practice of international relations in Europe from the 16th century onward. The balance of power is generally regarded as an essentially Eurocentric concept, although this assessment remains somewhat speculative because so little is known about the theory underpinning how other international societies have operated.

While the balance of power is generally regarded as a concept that is quintessentially related to international relations, it is also a concept that is regularly used across the social sciences as well as in general parlance. However, outside international relations, the balance of power is not seen to be particularly controversial. The term is employed ubiquitously because it acts as a metaphor that effectively transforms our conventional understanding of power. Unadorned, power is generally treated as an attribute that enables an agent possessing power to regulate the activities of other actors. In the literature on power, this usage is often referred to as a behavioral or agential conception of power. The balance of power metaphor transforms this conventional meaning into a structural or systemic conception of power.

When treated as a metaphor, the balance of power has been associated with a wide array of images, from a chandelier to the arch of a bridge. The image that most frequently springs to mind is a set of scales. From the perspective of international relations, the scales can be viewed as a system in which the weights that are placed on the pans of the scales are equated with the power capabilities of the states that operate within the system. What the metaphor demonstrates is that if the weight/power is increased in one pan, then the weight/power on the other pan is automatically affected. It follows that power must be treated as a relative rather than an absolute phenomenon. But the metaphor can also indicate that if there is a disproportionately heavy weight placed in one pan of the scales, then it is possible to establish an effective counterbalance by placing a number of lighter weights on the other pan. The even distribution of power that is then formed is often represented as the definition of a balance of power.

If the metaphor is brought into focus, however, it is clear that this very specific meaning should be subsumed under a much broader conception that associates the balance of power with a system in which power is a structural attribute. Such a system is, in effect, constituted by the distribution of power that forms within the boundaries of the system. So in a multipolar system, for example, there are several main poles of power, whereas in a bipolar system there are two main poles of power, and in a unipolar system there is only one main pole of power. This structural conception of power has attracted the attention of many historians, theorists, and practitioners from the past and the present who have wanted to understand international relations, and it has led them to presuppose that the behavior of states is or at any rate should be significantly influenced by the structure or balance of power that helps constitute the international system.

In general usage, by contrast, the same kind of assumption is not made. If a wife goes out to work and earns an independent income, it can be argued that the balance of power within the marriage will shift as a consequence. Or, in a multiparty system, if two opposition parties ally, then the balance of power within the political arena will alter. In both cases, the metaphor brings the structure of the system into focus. But there is no presumption that a change in the balance of power will necessarily strengthen or weaken the system. The metaphor simply highlights the change in the distribution of power. But in international relations, there is a long-standing political myth that states will respond to changes in the distribution of power in a way that ensures that international stability is restored. But the veracity of this myth has also long been challenged, and a significant countermyth has arisen asserting that the preoccupation with the distribution of power in international relations is a major source of conflict and disorder. This controversy about the balance of power has been carried forward into the contemporary discipline and accounts for the very distinctive role that the concept plays in international relations.

Historical Development of the Idea

Despite its putative importance, we still do not have a very developed understanding of how the ideas associated with the balance of power have evolved across time. Nevertheless, it is generally accepted that the contemporary conception of the balance of power can be traced back to the Italian city-state system and that the key ideas relating to the concept began to take shape in the 15th century. By the start of the 16th century, Francesco Guicciardini (1483–1540), a practicing diplomat and an acquaintance of Niccolò Machiavelli, was able to call on these ideas when he came to write *The History of Italy*. He wrote this history in response to the French invasion of Italy in 1494; his aim was to trace what he viewed as the tragic loss of independence by the Italian city-states in the first decades of the 16th century. He shows how these states failed to respond adequately to external pressures until eventually most of them were absorbed by the Hapsburg Empire. The book provides the first case study of states failing to follow the logic of a balance of power system and, as a consequence, succumbing to hegemony.

Guicciardini's use of balancing metaphors is restricted to the start of the book where he examines relations among the city-states in the final decades of the 15th century. The focus is on Venice, Naples, Florence, Milan, and Rome, and he acknowledges that relations among these city-states were characterized by suspicion and jealousy. But he also notes that Venice was by far the most powerful state in the system and, moreover, that it had hegemonic ambitions. But the logic of this situation as depicted by Guicciardini is that it preserved a balance of power system because, by forming an alliance, the other city-states were able to counterbalance Venetian power and thereby preserve the independence of all the city-states. Guicciardini, however, also acknowledged that rivalry among the less powerful states could potentially threaten the overall stability of the system.

But this stability was preserved, according to Guicciardini, because the ruler of Florence, Lorenzo de' Medici, recognized the importance of maintaining peace among the less powerful states and preserving the alliance against Venice. As a consequence, he was willing to operate as a balancer within the system. In other words, Guicciardini argues that Lorenzo had to constantly shift his own position in relation to the other states in the system in order to ensure that the overall balance was sustained.

What began to destabilize the system at the end of the 15th century was an emerging tendency by the city-states to invite outside states to help settle disputes among the city-states. The problem with this tactic, as Guicciardini saw it, was that states such as France outweighed the combined strength of all the Italian city-states. Nevertheless, when the Milanese concluded that Florence and Naples were forming an alliance against them, they called on France for assistance. But this move had the effect of destroying the Italian balance of power system and Guicciardini resorted to medical metaphors to characterize the subsequent developments. Eschewing the balance of power metaphor, he argued that external intervention was a medicine with effects that were far worse than the original disease.

The History of Italy was translated into other European languages, and the ideas associated with the balance of power were rapidly diffused across Europe during the 16th century. By the end of the century, therefore, the balance of power was viewed as a Europe-wide phenomenon. But during this period, the meaning of the balance of power was also extended beyond the idea of shifting alliances in a competitive or adversarial system. For millennia, there has been a metaphorical link between scales and justice, and during the 16th century, there was a growing tendency to discuss the balance of power in terms of a just equilibrium. In other words, the balance of power began to be associated with the establishment of a distribution of power that was not only regarded as stable but also fair. Initially it was argued that the balancer state must have this capacity to establish the link between stability and justice. But, over time, a stable and just balance of power was seen to arise from a general agreement among the great powers in Europe, and so the concept began to be viewed as a product of great power cooperation and mutual association. By the start of the 18th century, Europe-wide peace agreements such as the Treaty of Utrecht in 1713 began to make reference to a just equilibrium in their formal provisions. At this juncture, therefore, the balance of power was seen to provide the constitutional basis for what was called the Republic of Europe. As a consequence, it then became possible to extend the idea of the balance of power back to the idea developed by Greek and Roman theorists of a mixed or balanced constitution. But the relationship between these sets of ideas has yet to be teased out in any detail.

By the 18th century, however, there was also growing resistance to the idea that the balance of power could be treated as either some kind of a natural or a man-made law that could promote peace and stability. On the contrary, the balance of power was seen to be irrevocably tied to a system where war was endemic. By the 19th century, critics such as the British radicals Richard Cobden and John Bright were going further and arguing that the balance of power was an utterly meaningless concept. From their perspective, only by moving beyond the balance of power thinking was it possible to promote a peaceful world. The battle between advocates and critics of the balance of power persisted throughout the 20th century and, almost inevitably, when the study of international relations developed as a formal academic discipline during the present era, the balance of power immediately became one of its central but also deeply contentious concepts.

Contemporary Study of the Balance of Power

It is not difficult to demonstrate that the balance of power occupies a central position in the contemporary study of international relations because the concept lies at the heart of two of the very most influential theoretical texts published since the end of World War II. The first, Hans J. Morgenthau's *Politics Among Nations*, published initially in 1948, epitomizes classical realism and the second, Kenneth N. Waltz's *Theory of International Politics*, published in 1979, spearheaded what is now known as neorealism or structural realism. Despite the fact that these texts are still enormously influential, both texts and, in particular, their use of the balance of power have come under continuous criticism ever since they were published. It is certainly not the case, therefore, that either realism or the balance of power has ever occupied a hegemonic position in the field. Instead, it is more appropriate to see them as occupying the central ground but at the same time being engaged in a constant debate with critics coming at them from a number of very different directions.

As a classical realist, Morgenthau was well aware that he was working in the context of a long-standing tradition of European thought that embraced the concept of the balance of power, and he recognized the need to take into account both its adversarial and its associational dimensions. From his perspective, the American Constitution and other similar domestic arrangements provided the most effective examples of a functioning balance of power system, and it followed that, in an international context, the balance of power could work at best only imperfectly. For Morgenthau, the golden age for the balance of power was the era of dynastic international politics in the 18th century when states were governed by an international aristocracy that formed a distinctive international society. Under these circumstances, it was possible to achieve a just equilibrium by means of largely consensual and Europe-wide peace agreements. But even when these agreements broke down, Morgenthau argued that the uncertainties associated with any attempt to calculate the prevailing balance of power generated restraint on the part of the rival dynasties and encouraged them to operate on the basis of the established dynastic rules of the game. At the same time, however, he readily acknowledged that the world had moved a very long way from these conditions. In the aftermath of World War II, what he saw were two ideologically driven behemoths operating in the absence of any international society. The United States and the Soviet Union were restrained only by the crudest kind of balance of power, and he feared for the future of the world. Morgenthau argued that only if the two sides adopted the tools of classical diplomacy was there any hope that humans would survive.

Waltz approached the balance of power from a very different direction, and he reached a very different set of conclusions from those of Morgenthau. He was writing in the context of a detente that marked a period of relaxation in the Cold War tensions between the United States and the Soviet Union; however, his aim was to establish a general theory of international relations that would transcend historical difference and help reveal the essential differences between domestic and international politics and explain why the structure of the international system has proved so enduring.

In the first instance, therefore, Waltz argues that all political systems can be categorized under one of two headings: hierarchy or anarchy. In a hierarchical political system, actors are functionally differentiated and power is distributed on a vertical plane, so that actors can exercise power over subordinate actors in the hierarchy but will themselves be subject to the power of actors that occupy a superior position in the hierarchy. Having made the distinction, however, Waltz focuses his attention almost exclusively on anarchic political systems where actors are seen to operate on a horizontal plane. In other words, these actors do not consider that they are operating in a hierarchy or in a functionally differentiated system. They view themselves as independent and autonomous actors operating on the basis of self-help and so they are primarily concerned, in the first instance, with maintaining their independence and autonomy. To be able to do this, they must, from the start, establish how power is structured in the system. This requires them to identify the dominant actors in the system—those actors that possess a disproportionate amount of the overall power in the system. Cutting through complexity, Waltz distinguishes between bipolar and multipolar systems. Then, contrary to Morgenthau, he aims to show theoretically why multipolar systems are prone to generate an unstable balance of power, whereas bipolar systems are prone to generate a stable balance of power.

Focusing first on multipolar systems, Waltz argues that if the balance of power begins to move against any of the dominant states in such a system, they will be pushed to respond by either internal or external balancing. Internal balancing requires the actor to enhance its power position by domestic means. The most obvious mechanisms are either to expand existing military resources or to seek technological improvements. But Waltz also acknowledges that in a self-help or anarchic political system, there will be a tendency for any action that gives an actor a potential or future power advantage to be emulated by the other dominant actors in the system. As a consequence, he argues that in an anarchy there will be a tendency for actors to take the form of like units.

In a multipolar system, however, Waltz argues that when the balance of power shifts, there is the alternative strategy of external balancing. In other words, states can form alliances with each other to enhance their security. But Waltz is very well aware that alliances are an inherently unstable feature of any anarchic system. The literature on alliances reveals that an alliance generates at least two contradictory fears. On the one hand, there is the fear of entrapment, as the result of being drawn by an ally into an unnecessary or dangerous conflict. But, on the other hand, there is also the fear of abandonment. Such problems are endemic in a multipolar system. The problem persists even in the face of a rising hegemon when, rather than forming an overwhelming alliance, states will often choose to pass the buck and, at least in the first instance, leave it to other states to confront the hegemon.

Because of these sorts of uncertainties, it becomes extraordinarily difficult in an anarchic system to identify the nature of the balance of power at any point in time. Nevertheless, despite these uncertainties, Waltz argues that there is sufficient flexibility in the system to ensure that any potential hegemonic state will eventually be met by an effective counterbalancing alliance. But, by the same token, Waltz insists that it is very much easier to identify the state of the balance of power in a bipolar system: There are no alliances to complicate the assessment, and the two dominant states simply have to monitor each other's activities. As a consequence, not only are there fewer sources of instability, but it is also easier for the two dominant states to reach mutual agreement and thereby to move from an adversarial balance of power through to an associational balance of power, although Waltz does not use this terminology.

Waltz has proved to be extremely influential because he articulated a balance of power theory in such unequivocal terms, and it has encouraged other theorists to either revise or reject the theory. Offensive realists argue that dominant states aim to maximize their power position, and they characterize Waltz's approach as defensive realism because of the assumption that dominant states only aim to preserve their independence. They dismiss the idea of an associative balance of power and are much more willing to entertain the possibility of hegemonic success. In the same vein, theorists working from a world-historical perspective insist that anarchic systems transform into hierarchical or at least unipolar systems on a very regular basis. This poses a very significant problem for Waltz and for the balance of power theory more generally. But a more immediate problem for Waltz is the persistence of unipolarity in the post–Cold War world. Waltz insisted after the demise of the Soviet Union that unipolarity is a very unstable structure and that it would rapidly give way to multipolarity. But realist critics argue that there are few signs that any states are willing to compete with the United States in the military arena and that in any event Waltz's own Neorealist logic can be used to show why unipolarity is a very stable structure. It is unlikely that these debates are going to be easily resolved, and, as a consequence, the balance of power will continue to provide a theoretical focal point for theorists and practitioners in the future.

Richard Little
University of Bristol
Bristol, United Kingdom

See also Alliances; Anarchy; Bipolarity and Multipolarity; International Society; International System; Power and International Politics; Realism in International Relations

Further Readings

Gulick, E. V. (1967). *Europe's classical balance of power*. New York: W. W. Norton. (Original work published 1955)

Ikenberry, J. G. (Ed.). (2003). *America unrivalled: The future of the balance of power*. Ithaca, NY: Cornell University Press.

Kaufman, S. J., Little, R., & Wohlforth, W. C. (2007). *The balance of power in world history*. Basingstoke, UK: Palgrave Macmillan.

Little, R. (2007). *The balance of power in international relations: Metaphors, myths and models*. Cambridge, UK: Cambridge University Press.

Morgenthau, H. J. (with Thompson, K. W., & Clinton, D.). (2005). *Politics among nations: The struggle for power and peace* (7th ed.). New York: McGraw-Hill.
Sheehan, M. (1996). *The balance of power: History and theory*. London: Routledge.
Waltz, K. N. (1979). *Theory of international politics*. Reading, PA: Addison-Wesley.

Bargaining

Bargaining is an everyday activity that pervades all social life and, arguably, is the most important mode by which political decisions are made. It summarizes all social activities in which individuals or corporate actors have a common interest in working together but disagree on how they should cooperate. Bargaining is therefore a voluntary process through which negotiators try to distribute among themselves a mutual benefit. In politics, the mixture between cooperation and conflict that characterizes bargaining situations becomes apparent at all levels of decision making. An example of a recurrent, difficult bargaining process is the distribution of the taxed income of a polity between competing interest groups who advance contradicting claims to bolster their demand for a large share. Below, major game theoretical solutions to bargaining problems and their possible applications in political science are discussed.

In the simplest political bargaining problem, two negotiators disagree over how to divide a pie worth one unit of a fictitious currency (eurodollar). Any outcome in the "bargaining zone," which is delimited by the minimal individual payoffs the actors expect, can result from the negotiations. Trivially, no bargain is feasible if the minimal expectations do not overlap or, in other words, if the set containing the possible outcomes is empty. If the minimal expectation—technically often called the "reservation price"—of one player is zero, an extreme solution in which one side cashes in 1 eurodollar while the other side leaves the negotiation table empty-handed can occur, since the unlucky negotiator is assumed to be indifferent to the lack of a reward and the failure of negotiations. The "fair solution," in the absence of any power imbalance or neediness of a negotiator, is a 50:50 deal that splits the eurodollar into equal shares.

In many bargaining situations, actors demand a disproportional division of the spoils. The other negotiators then have to judge whether a credible claim that negotiations would fail without a concession from the other side supports the request. If they reject this demand, bargaining stops or is interrupted. Another suboptimal outcome materializes when negotiators accept a disproportional division and fail to call the bluff by the other side that the negotiations would break down if no concession were granted. The fate of a claim depends on the ability of the negotiators to communicate credibly that their claims are sincere. This creates in return an incentive to bolster the credibility of a specific position through strategic moves.

Social scientists who try to understand bargaining belong to either one of two camps. The first approach stems largely from social psychology and is, by and large, empirical. It tries to assess the importance of cognitive failures and other psychological features based on the performance of individual negotiators. In political science, the strategic theory of bargaining, as it was developed by Francis Edgeworth, Arthur Bowley, Frederik Zeuthen, and John Nash, is more influential as it takes into account that negotiation outcomes result from the interaction of at least two forward-looking actors. Technically, the strategic theory of bargaining is based on game theory. Its applications pertain to any subfield within political science and cover a wide variety of issues, ranging from the analysis of interstate war to delegation problems in public bureaucracies.

Bargaining games are answers to what is known as the "bargaining problem": Is it possible to predict the equilibrium outcome of a social negotiation in which all participating actors expect an individual gain? It took the genius of Nobel Prize winner John F. Nash to predict a unique equilibrium—that is, a single point—to the bargaining problem. In the Nash bargaining solution (NBS), this outcome represents the maximal product of the differences between the utility a negotiator attaches to his or her share x and the utility of the reservation price q.

By way of illustration, imagine a bilateral bargain in which the reservation price of two equally skilful negotiators is 0 so that actor i receives x and its counterpart j receives $1 - x = y$. We obtain the equilibrium prediction through the maximization of the Nash product $x(1 - x)$. Setting the differenced equation to zero ($1 - 2x = 0$), $x = 1/2$ and $y = 1/2$ result. A power-free bargain situation entices, in other words, a fair division.

As bargaining power has many facets, it can enter the calculus of the social outcome in various ways. One possibility is to equate power with capabilities c_i that enter as an exponential weight the utility function of the individual actor, where the sum of the capabilities add up to 1 within a society. Hence, if negotiator i is twice as powerful as counterpart j, the maximization problem in the bilateral case reduces to $x^2(1 - x)$. Differentiating $x^2 - x^3$ and setting the result to zero, we receive $2x - 3x^2 = 0$. If we disregard the outcome $x = 0$, actor i's share amounts to 2/3.

A further aspect of power at the center of the strategic theory of bargaining stems from the attractiveness of a so-called outside option. This outcome is equivalent to what an individual negotiator could reach if the negotiations fail or if they are stalled for some time. Another Nobel Prize winner, Thomas Schelling, made this point through the counterintuitive "paradox of weakness," according to which the less flexible negotiator possesses a bargaining advantage. Generally, the presence of an attractive outside option helps negotiators in their attempt to commit themselves credibly to a more beneficial bargaining position. Robert D. Putnam has used this insight in his "two-level games" metaphor, which stands for negotiations between governments that have to care about domestic pivot players. In such negotiations, governments who face a strong, isolationist opposition and a supramajoritarian ratification hurdle back home can convince a more cooperation-minded negotiation partner that only a minimal agreement is feasible. The commitment of such constrained governments is more credible precisely because they cannot get an agreement ratified that does not find the support of certain sections of the opposition. A treaty that is too close to the bargaining position of the foreign government is consequently doomed to fail in the negotiations or at the ratification stage.

The "Schelling conjecture" that the seemingly weak is in real life the strong actor can be illustrated within the NBS framework through the assumption that one of the actors can be committed to a disagreement point of 0.5 in the bargain over a pie of unit size. This credible claim lets the "zone of agreement" shrink and alters the maximization problem to $(x - 0.5)(1 - x)$, which leads after differentiation to a division of $x = 0.75$ and $y = 0.25$.

Criticisms of the NBS deal with its static nature, its axiomatic foundations, and its reliance on cooperative game theory. The first objection answers to the observation that most (but not all) negotiations consist of lengthy haggling. The second and the third of the perceived problems respond to the general criteria (axioms) that a bargaining solution has, in Nash's view, to fulfill in order to qualify as a socially rational outcome of a negotiation. His model, in particular, assumes in line with cooperative game theory that the negotiators agree on certain rules of the game and do not try to outsmart each other through unilateral moves.

A response to these objections is the development of noncooperative bargaining models that are dynamic and include relevant one-sided actions. Noncooperative bargaining models also often study the impact of information deficits on the bargaining outcomes. The most prominent analytical framework that assumes, in its most basic form, fully informed actors is the Ståhl-Rubinstein bargaining model. This sequential game introduces a strict bargaining *protocol* (a term standing for the order in which the actors are allowed to move) where the right to make an offer moves back and forth between two players. One version of this analytical framework assumes that actors discount the future with the factor δ_i ($0 < \delta_i \leq 1$). Hence, the longer the bargaining goes on, the less attractive it becomes. The model predicts again a unique outcome that is largely driven by the size of the discount parameters and hence by the patience of the negotiators. In general, the more patient an actor is, the larger is the share of the pie that she or he can bring home. If the intervals between the negotiation rounds converge toward zero, the Ståhl-Rubinstein bargaining model coincides with NBS.

The Ståhl-Rubinstein framework of analysis has been extended to negotiations among n actors. To make sensible equilibrium predictions in such a context, some additional assumptions about the bargaining protocol and the feasibility of certain strategies are necessary. The Baron-Ferejohn model of legislative bargaining, for instance, assumes that one member of parliament is randomly recognized as an agenda setter who can make equilibrium proposals to a majority of other legislators.

Generalizations of Schelling's insight that credible commitments are a key prerequisite to bargaining success incorporate models in which some negotiators are assumed to be imperfectly informed. Such "asymmetric information" pertains in the bargaining context most often to the credibility of a negotiator's claim that an outside option is more attractive than the proposal made by the badly informed negotiator. The better informed side tries in such a situation to improve the credibility of its claim through what is called "costly signals"; this is why such bargaining games belong to the category of "signaling games." In an interstate crisis, for instance, a costly signal can consist of a public warning or the mobilization of the troops. Signaling games have been successfully developed to understand conventional and nuclear deterrence; crisis behavior of this sort belongs to what is known as tacit bargaining, where actors use nonverbal means of communication to bolster their claims for a specific division of the spoils. The disadvantage of limited-information models, as games with asymmetric information are also called, is the plethora of possible equilibrium solutions. Oddly, this oversupply of predictions can only be trimmed down at the cost of asking the model agents to behave even more rationally.

The empirical evidence in favor of the equilibrium predictions that can be derived from bargaining models is mixed. In comparison with other decision-making modes such as voting or delegation, bargaining models predict outcomes more accurately. However, the extreme 100/0 division that some bargaining games suggest as a rational strategy hardly finds any support. Both laboratory and field experiments have shown that privileged negotiators often propose fair divisions against their own self-interest. The theoretical merit of these findings remains, however, controversial, although some powerful causal mechanisms that account for such seemingly irrational behavior have been proposed.

Gerald Schneider
University of Konstanz
Konstanz, Germany

See also Experiments, Laboratory; Game Theory; Rational Choice

Further Readings

Muthoo, A. (1999). *Bargaining theory with applications.* Cambridge, UK: Cambridge University Press.

Muthoo, A. (2000). A non-technical introduction to bargaining theory. *World Economics, 1,* 145–166.

Schneider, G., Finke, D., & Bailer, S. (2010). Bargaining power in the European Union: An evaluation of competing game-theoretic models. *Political Studies, 58*(1), 85–103.

Bayesian Analysis

See Statistical Inference, Classical and Bayesian

Behavioralism

The behavioral approach may be characterized as an attempt to conduct political science according to the example of the natural sciences—that is to describe, explain, and predict political phenomena as exactly as possible. It is the goal of behavioralism to gain scientifically valid—that is, methodically secured—generalizations on its subject. This method is not unique to the field of the social sciences; this application of the natural sciences to political science must be interpreted as being a part of a greater movement of which—despite great changes—all social sciences have become a part. From this point of view, behavioralism is a phenomenon of general social scientific empiricism that has found and still finds many supporters particularly in the United States. This entry reviews the beginnings of behavioralism and its evolution in the mid-20th century as a movement that reshaped the discipline of political science. It traces its emphasis on explanation, prediction, verifiability, and quantitative research and its growing influence in academic institutions as well as professional organizations. The entry concludes with a discussion of the goals and methodological assumptions that characterize the behavioral approach in political science.

Origins of Behavioralism

Although behavioralism as a trend began primarily after World War II, there are important predecessors in the context of American political science that paved its way. Beginning in the 1920s, the Chicago School and the New Science of Politics movement urged a greater use of empirical research to establish a scientific identity to the discipline. Charles E. Merriam, a leading figure in political science during those years, and Harold D. Lasswell, with his decades of work in the interdisciplinary area of policy studies, were key figures in those times; without their work, political science may well have taken a different course. Early behavioralist considerations developed in the 1920s as a reaction to the then predominant institutionalist approach, which was called "institutional realism," "realistic institutionalism," or simply "realism." The realist approach focuses on the analysis of the *actual* relations (in contrast to relations as legally defined) between different government institutions, political parties, and interest groups. In general, the traditional political-scientific analysis of political institutions aimed less at the development of theories than at finding facts. A strict distinction between factual claims and value judgments was an exception rather than the rule.

Originally, behavioralism was a protest movement due to dissatisfaction with traditional political science. The latter—in contrast to its neighboring sciences of sociology and economics—was accused of being characterized neither by cumulative research nor by reasonable scientific communication or even scientific cooperation. This flaw, it was stated, was among others due to political scientists having no common language, no common issues, and no agreement on methods. This was explained by the lack of comprehensive, generally accepted models or terminology schemes of the political process by which research was guided (David Easton, 1969).

This protest was initiated primarily by younger scientists who came from very different intellectual approaches and who were initially rather isolated, since they worked in different fields of political science. Protest was nourished by the experience of increased cooperation between political scientists and government institutions following the New Deal and, most of all, during World War II, when contact with representatives of the other social sciences had also increased. One result of this increased interaction between political scientists and politicians was the painful insight that there was an almost insurmountable difference between what was needed for political advice and what political science could offer. This difference between theory and praxis was not as strong in the neighboring disciplines, particularly psychology and economics, and, to a lesser extent, in sociology. In any case, government institutions were much less interested in advice from political scientists than from economists, psychologists, and sociologists, whose theoretical orientation and methodological basis were generally held in higher esteem than that of political scientists. In this context, the inability of traditional political science to predict events was particularly criticized.

The inability of traditional political science to predict and explain the rise of fascism in Europe was another reason for the dissatisfaction of behavioral-oriented young scientists. But fascism, particularly national socialism, also influenced the rise of behavioralism in another way: in the subliminal reorientation of the philosophy of science in American social sciences, influenced by emigrated European researchers. The popularization of Max Weber by Talcott Parsons and the publication of some of Weber's science-theoretical essays by Hans Gerth and C. Wright Mills (1946) were significant because they left out some of the pragmatist philosophical traditions. With many later behavioralists, the pragmatic philosophy, stating that value statements can be validated by practice, which was then predominant in the United States, was replaced by Weber's position that value statements cannot be proven by empirical means.

Another reason for the rise of behavioralism seems to be the changed international situation after 1945 and particularly the leading political role played by the United States in the postwar period. Increasing international interdependence, decolonization, and finally the Cold War confronted the political-scientific disciplines of international politics and comparative politics with tasks that could not be solved using traditional concepts such as diplomatic history, international law, and the comparison of institutions. Dissatisfaction with traditional political science was strongly supported by the exemplary effect of some social-psychological and sociological publications such as *The People's Choice* (1944), *The American Soldier* (1949), and *The Authoritarian Personality* (1950), where the empirical methods of random sampling, the questionnaire, and attitude analysis or panel analysis were employed and techniques of statistical evaluation were used to answer scientific questions. These methods previously seemed to have been reserved for economics and psychological diagnostics. These studies somehow served as signposts for a successful political-scientific behavioral research as it was imagined by early behavioralists.

It was predominantly the younger political scientists who in the early 1950s became fierce critics of the realist approach, most of all Easton. In several programmatic publications, reaching back as far as the early 1950s, he criticized the theoretical deficit of realist research, the lack of methodological reflection and terminological clarity, the absence of modern methods of data collection and analysis, and the one-sided emphasis on institutional aspects.

Realigning the Discipline of Political Science

In the beginning, the supporters of behavioralism were a small, scattered group of lonely fighters who were not bound together by any kind of organization and who felt connected to each other only by their way of understanding science. They met at seminars and in the context of conventions of the American Political Science Association (APSA) to discuss new methods of empirical research in politics. It was their goal to readjust the entire discipline of political science according to the criteria of behavioralism, not simply to add another subdiscipline. It is not surprising that such a demand for a fundamental readjustment of political science was strictly opposed by many established political scientists who met the new methods and theoretical perspectives with a lack of appreciation and even reacted with open dislike. As Dwight Waldo (1975) and Jürgen W. Falter (2001) have described, this resulted in the first phase of the so-called behavioralism controversy.

Behavioralism has been characterized by its interpreters and proponents as a "movement," as a "mood," a "conviction," and a "protest," as well as a "revolution" and a "renaissance." Easton distinguished between an intellectual trend and an academic movement; the former, he stated, had many more supporters than the latter. However, he said that it was difficult to identify both the supporters of the movement and the practitioners of the intellectual trend, because membership criteria were vague and the borders of the behavioralist conviction were unclear. In particular, he stated that it was almost impossible to distinguish "true" members of the movement, fellow travelers, and occasional sympathizers from each other.

Easton argued that behavioralism did not show any formal organization as a movement. Although there had been short-term plans in this direction, efforts to bring the movement together had not been very fruitful. On the other hand, there had always been a feeling of being loosely connected and an agreement on basic assumptions and scientific ideals, as well as lively mutual communication. Also, certain spokespeople of the movement had always been accepted by the scientific community. The success of behavioralism, he stated, was due not only to its effective leaders, its capable supporters, and the exemplary effect of important publications but also to a number of favorable conditions. In addition to the funding policies of the great foundations, the significance of the fact that the Social Science Research Council and its committees provided the behavioral movement with organizational vehicles for the development of American political science cannot be overemphasized.

Growing Institutional Influence

In the 1950s and the early 1960s, the increase in numbers of supporters of the behavioral approach seemed unstoppable. It experienced early, important success, as Harold D. Lasswell, perhaps the most important intellectual pioneer of behavioralism, was elected president of the APSA in 1955–1956. His successor, V. O. Key Jr., who was president in 1957–1958, was a behavioralist as well. Under their presidencies, for the first time, the APSA established a special team dealing with questions of political behavior. The number of behavioral-oriented teams at APSA meetings grew considerably until 1960. The rise of behavioralism became manifest in a dramatic rise of research papers after 1960. Soon there was no longer any subfield of political science without behavioral-oriented studies, although naturally the extent to which behavioral thought penetrated the various subfields varied greatly. The focus of research was the analysis of decision-making processes and patterns of behavior. The individual served as a basis for collecting data, and the individual actor, the group, the institution, or even the political system in its entirety, was the subject on which statements were made. The core concepts of behavioral research took shape mainly in the work by Herbert Simon on political decision making, Heinz Eulau on roles, David Truman on interest groups, David Easton on the political system, Harold Lasswell and Abraham Kaplan on the concept of power, and others. Political-scientific research now focused on theoretical concepts rather than on the analysis of everyday politics as the realists had preferred.

From about 1960 on, a number of important research papers based on the model of behavioralism were published, among them *The American Voter* by Angus Campbell and Philip E. Converse (1960), *The Civic Culture* by Gabriel Almond and Sidney Verba (1963), and Karl Wolfgang Deutsch's *The Nerves of Government* (published in German in 1963). In addition to studies on elections developed at Columbia University and the University of Michigan (*Voting* and *The Voter Decides*), several programmatic and empirical studies were published that were targeted at fundamentally restructuring the subfields of comparative politics and international relations. After 1956, the profile of the *American Political Science Review* also changed drastically, indicating a gradual change in the discipline. Clearly, the quantitative element was gaining ground. Now, almost all American universities appointed behavioralists to their departments of political science, which triggered off a kind of chain reaction, further supporting the spread of the behavioral conviction. Studies written in the spirit of behavioralism started to dominate the appearance of the most important scientific magazines in the United States. With just a few exceptions, for a number of years after 1964, almost all APSA presidents came from the behavioral movement or were at least related to it. In the course of only 1½ decades, behavioralism had succeeded in fundamentally changing political science in the United States.

In the course of the second phase of the struggle of the behavioral research program, following the students' revolt of the 1960s, a strongly application-oriented dimension was added to behavioral political science. In this, policy studies tried to combine a strictly scientific way of proceeding with practical relevance in order to provide political advice. Although it has often been declared dead, the behavioral research program is still alive, and contemporary scientists working quantitatively and empirically would call themselves behavioralists.

There has always been a fundamental difference between empirical and theoretical behavioralism. Both were based on the same philosophy of science—that is, logical empiricism. However, they deal with different questions and use different theoretical analysis units. Whereas a great part of empirical research earlier referred to and still refers to explaining individual behavior, theory-oriented behavioralism was mainly interested in the political system. Both kinds of behavioralism turned against philosophically normative, purely institutionalist, or social criticism–oriented approaches of political science. As Falter (1982) notes, a bitter behavioralist struggle in American political science resulted, from which, however, the behavioral research discipline survived relatively unscathed.

Goals and Methodological Assumptions

In this section, the 10 most important goals and methodological assumptions of behavioralism are summarized and commented on. This catalog of behavioral principles tries to comprehensively explain the common core of the various behavioral trends in the context of political science, as far as it is possible to bring the variety of approaches together.

The Need for Theory

The goal of behavioralist political science is not merely to describe political processes but also to explain and predict them. In this context, theory is seen as guiding the selection of research topics and the description and integration of empirical findings.

Looking for Regularities

According to the behavioral basic assumption, one must reach back to law statements in order to be able to explain and predict political processes. One pragmatic precondition for the search for regularities is the assumption that social and political processes are subject to such laws. However, with its search for laws, behavioralism does not restrict itself to empirical generalizations but tries to go as far as possible to theoretical statements. This makes it different from the classical and radical behaviorism of John Watson and B. F. Skinner with which—erroneously—it is often equated. This, at first, only a verbal equation, which, however, is often supposed to disqualify behavioralism as being "positivist" and thus out-of-date, blurs important, fundamental differences between the two trends. Behavioralism, which indeed is also an analysis of behavior using empirical and quantifying methods, goes far beyond the purely descriptive recording and generalization of behavior. Psychological behavioralism, being a theory-guided empiricism, (a) looks primarily for explanations and not for empirical generalizations and (b) includes intervening variables moderating between stimulation and reaction, for example, cognitive consciousness processes, in the form of theoretical constructions, in its analysis. Thus, in contrast to purely empirical generalizations, the theoretical statements of behavioralism deal not only with processes that could be directly observed but also with dispositional factors such as personality traits and attitudes that by empirical means can only be described indirectly—that is, with the help of indicators. These latent dimensions, which cannot be directly observed, are described with the help of theoretical concepts—that is, by so-called constructs. An example of such a theoretical construct is the concept of identification with a political party. Theoretical constructs are connected by definitions and lawlike statements. These theoretical concepts are connected to the observation level by so-called correspondence rules, which make the empirical evaluation of a theory possible. According to the behavioral opinion, all theoretical concepts must be operationalized (or rather be operationalizable).

Striving for Verifiability and Objectivity

It is typical for science, in the behavioral sense, that all statements must be verifiable. From the empirical point of view, statements are verifiable only if they either immediately refer to observable facts or can at least be based on statements formulated in the observational language. In contrast to the analytical philosophy of science, which only demands that statements must be (at least indirectly) observable or that they can be falsified, behavioralism still follows the verification principle of neo-positivism. For it, sentences that cannot be verified must be excluded from the scientific canon. To meet the criterion of verifiability, the measurements of the empirical researcher must be as objective as possible—that is, independent of the individual scientist. Objective research results are characterized by the possibility of their being repeated by other scientists under the same circumstances. To be verifiable, the applied research tools must provide reliable measured values. A measurement tool is reliable if it produces the same results in the case of repeated measurement of the same events, with a certain leeway for mistakes that must be as small as possible. This way, however, one does not know if a reliable measurement tool, for example, a scale, really measures that what is supposed to be measured. For this reason, measurements must not only be formally reliable but also topically valid. In this context, the formal exactness of measurement is a necessary but not sufficient condition for the topical exactness of measurement. To guarantee the verifiability or more generally the meaningfulness of its results, empirical research must also argue in a way that is without logical contradictions, that explicitly defines its terms and uses them consistently, and that excludes all concepts from the scientific language that do not refer to empirical facts.

Elaborated Research Techniques

Empirical social research employs a number of tools for data collection and data evaluation. These include various methods of attitude measurement, techniques such as the interview, the methods of quantitative topical analysis, and standardized observation schemes. Among the most important research techniques of empirical political science are various random sample procedures. Statistical methods of evaluation include regression analysis, variance and covariance analysis, as well as factor and path analysis, among others. It is the goal of the statistical methods of evaluation to cope with extensive amounts of data, to determine connections between analyzed variables, and to determine the degree of safety by which the measured connections can be accepted. In this context, the use of constantly improved, refined research techniques is not an end in itself but a means to meet the previously described quality criteria of research.

The Trend Toward Quantification

Inevitably, the demand that measurement results must be as exact as possible results in the attempt to make data collection and data evaluation as precise and standardized as possible. In this way, the tendency toward verifiability and comparability of statements is increased. According to its own statement, the call for quantification is not meant as a dogmatic demand but as a useful goal to strive for where it is reasonable for the object. Other goals, such as the theoretical and practical relevance of research results, are not supposed to retreat. Often—and sometimes with justification—the opponents of behavioralism object that its most important characteristic is its quest for maximum quantification while often losing focus on the subject under analysis.

Focusing on Individual Behavior

According to the behavioral point of view, all statements must be referred to as observable behavior. On rare occasions, the behavioral concept of behavior also refers to small groups. However, the behavioral concept of behavior is very comprehensive and includes both acts of behavior in the stricter sense and the expression of attitudes, intended behavior, or values as they are collected by using questionnaires or scales. If at the beginning of the behavioral movement, its spokespeople assumed that empirical and theoretical analysis units could be made to coincide—that is, the levels of observation and statement match—soon this behavioral-psychological remnant was abandoned. Political science deals also with groups, associations, and institutions that cannot be excluded from the field of political science just to meet a methodological demand. Otherwise, behavioral political science would actually have been reduced to the social psychology of the political behavior of individuals, as is often claimed by its critics.

Inductivism

Another typical feature of behavioral research activity is its inductive way of proceeding. Usually, research happens by observing empirical regularities—for example, that Catholics on the European continent vote for conservative parties more often than Protestants do. These empirical regularities are conceptually expressed by way of law statements. Also, theoretical constructions are definitely used—for example, that of the dogmatic personality. According to the behavioral point of view, these theoretical constructions, which are employed to explain and predict political behavior, must always be directly operationalized, which, however, is not at all necessary according to the point of view of the analytical philosophy of science. Thus, they stay comparatively close to the empirical surface, which might be one of the reasons why behavioralist theses are comparatively less fruitful, which, however, in this respect is hardly different from the theories of its neighboring sciences of sociology and social psychology. What exist in rich quantities are empirical single results and culture-specific generalizations, which, however, are comparatively unconnected.

Value Relativism

The behavioral position and the philosophy of science share the opinion that it is not possible to give reasons for value judgments with the help of empirical methods alone. Only cognitive statements can be empirically verified—that is, statements on the world as it is or, rather, as it appears to the scientist, but not statements on what the world should be like or how the scientist would like it to be. Thus, according to the behavioral opinion, value judgments are not capable of truth: They cannot be inferred from statements on facts. According to Arnold Brecht, between *is* and *ought* there is a gap that cannot logically be bridged and that makes it impossible for the scientist to present value judgments as a result of his or her research. Value positions that are included into empirical statements will distort the results of research. Thus, the individual scientist must try to neutralize his or her value tendencies and prejudices. This is why empirical science follows the ideal of impartiality in this sense, although it is in some way like the "Blue Flower" for scientists—a goal that they will never fully achieve. Even according to behavioral opinion, the individual will always be somewhat a prisoner of his or her value preferences.

Giving up on empirically reasoned value judgments does not rule out that values might be empirically analyzed as dependent or independent variables. Furthermore, according to Immanuel Kant, empirical political science is capable of analyzing systems of value preferences in respect of their consequences as well as of showing inconsistencies and incompatibilities, of researching the conditions for their realization, and of giving technological recommendations in the form of hypothetical imperatives.

Orientation Toward Basic Knowledge

The use of scientific insights for the realization of societal goals is possible only if there is a sufficient amount of basic knowledge. But even today this is not the case with most subfields of political science. Even after approximately 60 years of behavioral-oriented research, sufficiently confirmed theories are rather an exception. On the other hand, application-oriented research contributes little to the growth of scientific basic knowledge. To be able to make sound predictions, according to the behavioral program, there must be knowledge of behavioral regularities as well as theories using general terms, among others. The technological application of scientific insights depends on the existence of a knowledge stock that must be worked out by way of basic scientific orientation. That is why (at least originally) behavioralism supports a program of "true science," although as a consequence of Easton's postbehavioral revolution we observe that, in this respect, points of view are not so strict anymore. Today, many younger researchers, who nevertheless are definitely connected to behavioralism in respect of their research techniques, try to combine theoretical and practical relevance, most of all in the field of policy studies.

Integration and Interdisciplinarity

For its efforts to work out a secured and growing knowledge stock, political science must reach back to the concepts, procedures, and results of other social sciences. According to the behavioral opinion, politics include only one field of the social; personality traits in the widest sense may be as significant for political behavior as social or economic facts. It is necessary to reach back to data collection and evaluation techniques of other social sciences because there are only a few empirical methods developed by political science itself. Almost all the research techniques used by it are takeovers or modifications of sociological, social-psychological, economic, or individual-psychological methods and give reason to the demand for an interdisciplinary approach and the integration of research results.

The Lasting Influence of Behavioralism

Even if few political scientists today consider themselves behavioralists in the stricter sense, it is a fact that at first American political science and later international political science have been characterized and penetrated by behavioralism to such an extent that in Thomas Kuhn's sense—with a grain of salt—we may speak of a scientific turnaround whose heuristic possibilities do not seem to be exhausted at all even after 60 years. It is an achievement of the behavioral revolt that today empirical-quantitative studies are a matter of course. As for the methodical-statistical refinement of analysis, for a long time empirical political science has been acting at the same level as its neighboring disciplines. The bitter quarrels between doctrines from the 1950s to the 1970s have been replaced by a peaceful coexistence of many approaches and theoretical positions. In so far, Robert Dahl proved to have almost prophetic qualities when he subtitled a 1961 article with an early obituary for behavioralism, "Epitaph for a Monument to a Successful Protest."

Jürgen W. Falter
Johannes Gutenberg-University Mainz
Mainz, Germany

See also Attitudes, Political; Electoral Behavior; Measurement; Policy Analysis; Political Culture; Positivism; Quantitative Methods, Basic Assumptions; Survey Research

Further Readings

Berelson, B. (1968). Behavioral sciences. In D. L. Sills (Ed.), *International encyclopedia of the social sciences* (pp. 21–45). New York: Macmillan.

Dahl, R. A. (1963). *Modern political analysis*. Englewood Cliffs, NJ: Prentice Hall.
Deutsch, K. W. (1963). *The nerves of government*. New York: Free Press.
Easton, D. (1969). The new revolution in political science. *American Political Science Review, 63,* 1051–1061.
Eulau, H. (1963). *The behavioral persuasion in politics*. New York: Random House.
Falter, J. W. (1982). *Der "positivismusstreit" in der Amerikanischen politikwissenschaft* [The "positivist dispute" in American political science]. Opladen, Germany: Westdeutscher Verlag.
Falter, J. W. (2001). Behavioralism: Political. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 1125–1128). Amsterdam: Elsevier.
Lasswell, H. G. (1948). *The analysis of political behavior*. London: Routledge & Kegan Paul.
Somit, A., & Tanenhaus, J. (1967). *The development of American political science*. Boston: Allyn & Bacon.
Waldo, D. W. (1975). Political science: Tradition, discipline, profession science, enterprise. In S. I. Greenstein & N. W. Polsby (Eds.), *Handbook of political science* (pp. 1–130). Reading, MA: Addison-Wesley.

BELIEFS

Political beliefs can be defined as individual psychological orientations toward objects of a cognitive nature in the political world (e.g., polities, institutions, actors) that consist of the perceived likelihood of a given attribute being attached to them. Therefore, a political belief has three basic components: an object, an attribute, and the perceived subjective probability that the object and the attribute go together. However, as with other concepts such as values or attitudes, political beliefs are hard to define properly, having been subject to different and even contradictory uses. Furthermore, any survey of the literature will show how political beliefs are measured in different and often imprecise ways. It is, thus, a concept still "open" to full discussion, or, as Imre Lakatos would have put it, it is a concept still in the "morphological phase." Therefore, this entry first reviews the main uses and definitions of the concept found in the literature to appropriately frame the definition advanced above. This concept of beliefs is then discussed in the wider context of belief systems.

Conceptual Problems

Generally, one can think of political beliefs as individual orientations toward politics, related to other similar concepts such as political values and political attitudes. But the terminological jungle that comes along with these concepts makes it very hard to make sense of such a definition; as William J. McGuire has stated, we find ourselves with a set of names in search of a distinction rather than with a distinction in search of a terminology, making the situation rather complicated and blurry. An often implicit assumption is that these concepts (beliefs, values, and attitudes) can be placed on an underlying continuum ranging from the most basic or abstract to the most specific. The most abstract ones would therefore be the filters through which citizens form their more specific orientations. However, this continuum is often defined by the opposition between attitudes and values. Values are commonly placed on the top of the scale, at the highest level of generality or abstraction, and attitudes are conceptualized as being more specific. But the place of political beliefs within this framework is not so clear. A review of the uses of the concept shows that there is deep controversy over how to locate political beliefs within this context. Indeed, on one hand are scholars who quasi automatically equate political beliefs with values and use both terms almost interchangeably, such as Stanley Feldman, who treats "core" beliefs as roughly equivalent to political values. On the other extreme, there are many scholars who consider beliefs as the more specific components of attitudes. In this latter framework, attitudes are conceptualized as summary evaluations made of opinions and beliefs, and therefore, beliefs would be placed at a lower level of generality or abstraction.

This controversy often passes unnoticed because most scholars simply do not pay attention to the definitional issue. Nevertheless, the extremely divergent uses of the concept must lead us to conclude that the placement of the concept in a latent (undefined) scale is not the appropriate path to follow if we aim at analytical precision. There is no clear consensus on where to place beliefs in such a scale and, therefore, its usefulness is extremely limited. We thus turn our attention to other attempts that are more analytically oriented.

These contributions are to be found especially in the field of social psychology. Martin Fishbein and Icek Ajzen advanced perhaps the clearest definition of *beliefs*: They define the term as the perceived likelihood, or the subjective probability, that an object has a specific attribute or characteristic associated with it. Therefore, political beliefs would be presented only by statements of the following nature: "The government is efficient," "The government is responsive," "The judiciary system is fair," or "Health care is free for all," and so on. These are cognitive thoughts and ideas that subjectively link specific attributes to political objects. Additionally, as stated by Oskar Niedermayer and Bettina Westle, beliefs can differ in strength, certainty, and subjective relevance or salience. Thus, they should be measured on a single dimension of subjective probability of an object–attribute relationship (i.e., how likely or certain it is that object *X* has the attribute *Y*), commonly through the degree of agreement or support to a given statement that must be both object and attribute specific.

This conceptual framework allows for a clearer conceptualization of the term, and it lays the ground for a distinction between beliefs and attitudes. Beliefs are primarily of a cognitive nature, while attitudes are evaluative. As Lester Milbrath states, beliefs are cognitions with a feeling of credulity attached to them. In other words, a belief contains a cognitive component of an object (it has to be known for the belief to exist) linked to "credible (believable) or incredible (unbelievable)" features of that specific object. In this sense, beliefs are distinct from values and attitudes, since the latter imply the attachment of a feeling of valence (like/dislike) to the cognition rather than to credulity. (In Milbrath's scheme, both beliefs and attitudes have a cognitive dimension since one can only believe or evaluate known objects.) Both feelings, Milbrath insists, are interrelated (i.e., we tend to believe what we like), but at the same time, they are different (we can believe things that we dislike and vice versa). In any case, the argument is the same: Beliefs are based on the perceived probability that an object has a given feature associated with it. An implication of this conceptualization is that beliefs can often be falsified with external, objective criteria, while this is not the case for attitudes and values.

However, as Stuart Oskamp notes, this way of conceptualizing beliefs and attitudes raises the problem of the status of the so-called evaluative beliefs that combine the perceived likelihood of having a given attribute on one hand, and an evaluation on the other (e.g., "The government is inefficient," "The authorities are responsive," or "The political system is democratic"). This is so because the attributes being believed or disbelieved have a clear and widely shared positive or negative implication. Even if these statements are syntactically nonevaluative, since they refer to the likelihood of an attribute being attached to an object, semantically they are clearly evaluative orientations. This ambivalence makes evaluative beliefs almost indistinguishable from attitudes. Evaluative beliefs might be considered as an intermediate category, but what they reveal is indeed a continuum that ranges from the most objective (and falsifiable) beliefs to the most evaluative ones.

The inability to distinguish beliefs from attitudes in this kind of literature points again to a serious flaw in conceptualization. As noted above, any meaningful definition of beliefs must, at least, allow us to separate them from political attitudes. Scholars using this theoretical framework try to resolve this problem by going back to the same kind of scale of generality or abstraction: Attitudes are considered to be generalized and enduring evaluations of an object, while beliefs are singular and specific statements (evaluative or not) that refer to particular characteristics of the object. Therefore, most scholars of this framework consider that a person's attitude toward an object can be thought of as a summary of his or her evaluative beliefs. For example, the overall evaluation of the political system would be an attitude, while the evaluation of each of its specific features (responsiveness, efficacy, transparency, etc.) would, within this framework, be thought of as a belief.

There is an additional empirical problem in this conceptualization: Cognitive beliefs are the foundations for evaluative attitudes and, therefore, we must expect a close empirical relationship among them. However, this relationship is not present in many instances. Scholars justify this lack of empirical evidence by saying that cognitive orientations and evaluative attitudes may not be empirically related. However, attitudes made up of summary evaluations must show—at least partially—*some* congruence with the set of beliefs that produces such general evaluations. Without such an empirical congruence, therefore, we should not think of those attitudes as summarizing a set of beliefs.

Ideology and Political Beliefs: The Belief-Systems Debate

Despite the relevance of the preceding definitional discussion, we must acknowledge that the core debate in political science on political beliefs has basically revolved around the concept of belief systems—their nature, structure, and internal consistency. This debate was originated by Philip Converse's seminal work *The Nature of Belief Systems in Mass Publics*. The following section briefly reviews the concept and the terms of the debate, which has been both conceptual and empirical.

The key concept is political belief systems. In Converse's framework, belief systems are coherent structures of "attitudes and ideas" whose components are logically organized. The internal coherence of belief systems stems from political ideology—mainly in terms of left-right ideology. Ideology, therefore, consists of an overarching set of integrated principles about politics and the social world from which specific beliefs are deductively derived. This concept was meant to overcome the limits of the term *ideology*, since its multiple (and often contradictory) uses and its normative overtones made it of limited utility for research. How do belief systems acquire their structure and eventually their internal coherence? Converse's perspective, which has its foundation in physics, implicitly assumes randomness to be, somehow, the natural condition of things. This underlying assumption implies that whenever we find a given structure in citizens' political beliefs, we must look for an external force, which, in his framework, is called a "constraint."

What are these external forces, or constraints? Converse discusses two types of constraints: the social and the psychological. However, his focus, and the focus of the literature following him, has been disproportionately placed on the first one. A social constraint for the political belief system comes from the dominant role of political elites, who "package" attitudes and beliefs in coherent structures for consumption by the public, as expressed by Stanley Feldman. These packages are often referred to as "political ideology." The psychological constraints are based on the role of superior convictions concerning humans and society that bind together more specific orientations, such as attitudes and beliefs. In other words, Converse is referring to what has otherwise been called "values" or, in Stanley Feldman's terminology, "core beliefs." Therefore, the ideological structure of mass belief systems (in terms of left-right opposition) might be thought of as a "social" or elite-driven constraint, while the alternative perspective of values and core beliefs, or even non-political normative judgments, may fall under an alternative psychological constraint.

The actual structure (or lack of structure) of citizen belief systems, as well as its main determinants, has also been the subject of an enduring controversy. The empirical findings to test this theory have displayed very little consistency across time and across issues or beliefs. Only a reduced segment of the publics (the so-called ideologues) shows a coherent pattern of beliefs congruent with the basic left-right ideological scheme. This low interitem congruence and overtime stability may signify a limited political understanding of citizens in ideological terms, but it could also be an artifact or measurement error. Furthermore, Converse himself has shown that party identification is substantially more stable over time than ideological self-positioning.

More recent research has led to new reviews of Converse's findings. For example, John Zaller has shown in *The Nature and Origins of Public Opinion* how citizens' beliefs and attitudes are changed and polarized according to political elites' discourses even within the group of citizens in which ideologies are not structured and coherent belief systems are not dominant. Additionally, comparative research has shown how the extent to which citizens hold ideologically structured belief systems also depends on some contextual features, such as the political party system, the structure of party competition, and the electoral system.

Mariano Torcal and Jordi Muñoz
Universitat Pompeu Fabra
Barcelona, Spain

See also Attitudes, Political; Ideology; Public Opinion; Values

Further Readings

Converse, P. E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 206–261). New York: Free Press.

Feldman, S. (1998). Structure and consistency in public opinion: The role of core beliefs and values. *American Journal of Political Science*, *32*(2), 416–440.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Milbrath, L. W. (1968). The nature of political beliefs and the relationship of the individual to the government. *American Behavioral Scientist*, *12*(2), 28–36.

Niedermayer, O., & Westle, B. (1995). A typology of orientations. In O. Niedermayer & R. Sinnot (Eds.), *Beliefs in government: Vol. 2. Public opinion and internationalized governance* (pp. 33–50). Oxford, UK: Oxford University Press.

Oskamp, S. (1977). *Attitudes and opinions*. Englewood Cliffs, NJ: Prentice Hall.

Rokeach, M. (1968). *Beliefs, attitudes and values: A theory of organization and change*. San Francisco: Jossey-Bass.

Zaller, J. (1992). *The nature and origins of public opinion*. Cambridge, UK: Cambridge University Press.

BILATERALISM

Bilateralism refers to any relationship between two parties. In the study of international relations, we normally think of bilateralism as referring to relations between two states. But this is an overly narrow definition. Bilateralism can denote an arrangement between two private companies—economists, for instance, write of a "bilateral monopoly" where there is effectively only one seller and one purchaser in a market. Alternatively, bilateralism can involve two nongovernmental organizations (NGOs), two intergovernmental organizations, or any two-party combination of these actors (including states). For instance, a bilateral agreement is negotiated when an intergovernmental organization, such as the United Nations Children's Fund (UNICEF), commissions a local NGO to manage a project for it.

The term *bilateralism* is also used to refer to a strategy of how relations are conducted; it denotes an approach that prioritizes reaching agreement with one other party rather than acting by oneself (unilateralism) or together with two or more other parties (multilateralism). It is also differentiated from regionalism, where relations are conducted with two or more other parties within the same geographical area. Bilateralism need not involve geographical proximity—an instance is the U.S. free trade agreement with Bahrain.

The vast majority of treaties between states are of a bilateral character. These cover the full range of diplomacy, including agreements on extradition, overseas development assistance, cultural and educational exchanges, aviation, double taxation, cooperation on criminal and terrorism issues, the exchange of nuclear materials, postal services, visa requirements, mutual acceptance of domestic standards, cooperation on meteorological services and space exploration, as well as defense, investment, and trade issues. It is in the areas of investment and particularly trade that strategies of bilateralism have become most controversial in recent years: The focus in the remainder of this entry is on these areas.

Bilateralism has been the dominant form of international trade agreement since the start of international commerce. Such treaties have a history as long as that of international trade itself—the first known bilateral trade agreement was between Egypt and Babylonia, dating to 2500 BCE. As international commerce expanded, so did the number of bilateral trade treaties multiply, driven first by the establishment of the great maritime empires of the 16th and 17th centuries and then by the Industrial Revolution in the 18th and 19th centuries.

These bilateral agreements customarily assumed the form of a treaty of friendship, commerce, and navigation, which typically outlined the terms under which trade and shipping between the two states would take place and the rights enjoyed by individuals and firms from one state who lived, conducted business, or owned property in the partner state. The treaty also detailed the tariff treatment to be applied to various goods.

In the 19th century, in particular, these bilateral commercial treaties were used not just to protect existing commerce or to establish, on a reciprocal basis, new markets for the two parties but also to impose decidedly unequal arrangements on militarily weaker states. The most notorious examples were the treaties imposed by Britain on Qing Dynasty China after the First Opium War (1839–1842). The Treaty of Nanking opened five ports to British exports, at fixed tariffs to be agreed between the two governments, and ceded Hong Kong to Britain. Other industrializing countries followed suit, the United States, for instance, negotiated a commercial treaty with China in 1884 that was markedly unequal in the concessions made by the two parties.

Although the development of the unconditional most-favored-nation principle by leading European trading states in the 19th century laid a foundation for multilateral commercial treaties, bilateralism remained the norm throughout the interwar period. The decline in commerce following World War I prompted an innovative attempt at multilateral cooperation on trade through the staging of a World Economic Conference in 1927, but the treaty negotiated failed to gain enough signatories to come into effect when countries retreated into protectionism as the depression deepened at the end of the decade. In response to the Great Depression, governments typically sought to negotiate bilateral commercial treaties.

Only with the establishment of the General Agreement on Tariffs and Trade (GATT) in 1947 did a multilateral approach to trade cooperation gain precedence over bilateral treaties. The GATT established rules intended to ensure that countries that agreed to participate in multilateral rounds of tariff reduction negotiations would not be able to act unilaterally to negate the obligations to which they committed. The GATT, however, did permit some exceptions to its rule—that contracting parties must unconditionally provide most-favored-nation status to other parties. It thereby opened the way for bilateral agreements to continue to play a role in the governance of global trade. First, it "grandfathered" existing arrangements between imperial powers and their current and former colonies. Second, it permitted the negotiation of free trade agreements and customs unions. Although the original expectation was that the latter provision would apply primarily to regional arrangements, it has increasingly been used to legitimate bilateral arrangements.

For most of the period between the establishment of GATT and that of its successor, the World Trade Organization, in 1995, multilateralism dominated the governance of international trade. All the major trading powers (the United States, the European Union [EU], and Japan) gave strong support to multilateral liberalization despite domestic resistance to the opening up of "sensitive" sectors. In the 1980s, however, bilateral strategies gained new prominence on two dimensions.

First, with the rapid growth of exports from East Asia that caused problems for domestic manufacturers in North America and Europe, and led to rapidly burgeoning trade imbalances, governments increasingly turned to bilateral agreements (as well as unilateral measures) to alleviate these problems. Most prominent were the actions taken by the U.S. government under Section 301 of the U.S. Trade Act. The 1962 Trade Expansion Act had contained a provision granting the U.S. president authority to impose retaliatory measures when foreign governments were considered to have harmed U.S. trade interests. This measure was substantially strengthened by Section 301 of the 1974 Trade Act, which moved the primary responsibility for action from the president to the U.S. Trade Representative (USTR).

Section 301 provided that in the event of the USTR determining that foreign trade practices were unfair, negotiations should take place, and, if these failed to realize a satisfactory outcome, the United States could then impose restrictions on its partner's exports or withdraw existing concessions. Although the act allowed for unilateral action, the outcome of the Section 301 investigations often took the form of bilateral negotiations. These produced agreements with a number of East Asian and other industrializing economies (most notably Japan, Korea, and Taiwan), under which their governments undertook to limit the volume of exports of specific products (at various times, automobiles, steel, textiles, clothing, and footwear). These arrangements circumvented GATT regulations in that the country responsible for exports "voluntarily" undertook to restrain them. In 1984, the EU implemented a commercial policy regulation that had similar provisions to the U.S. legislation. "Voluntary export restraints" were outlawed by the World Trade Organization (WTO) Treaty: Such bilateral agreements, subsequently, have largely disappeared.

Meanwhile, a new form of bilateralism gained prominence from the mid-1980s: preferential trade agreements (PTAs). As noted above, the GATT exempted parties to free trade areas and customs unions from the requirement of applying for most-favored-nation treatment. In the 1960s, a wave of PTAs had been negotiated by developing economies—but most of them proved ineffective. Among industrialized economies, only the European countries—through the EU and the European Free Trade Association—had pursued regional strategies that led to significant exceptions to the multilateralism of the GATT. The United States and Japan (and other East Asian economies) were enthusiastic supporters of multilateralism. Their attitude began to change—in part because of concerns about the "unfair" trading practices of East Asian economies and about a perceived loss of bargaining power vis-à-vis the EU. The key development that symbolized a U.S. move away from a multilateralism-only strategy was the negotiation of its first bilateral free trade agreement—with Israel—in 1985 (the United States had earlier entered into a bilateral sectoral agreement with Canada—the U.S.–Canada Automotive Products Agreement of 1965—and had introduced a system of trade preferences for Caribbean countries in 1983). The Canada–U.S. Free Trade Agreement followed in 1989; in turn, this was converted into the North American Free Trade Agreement in 1993 with its extension to Mexico.

Other countries quickly followed suit; the consequence was a veritable explosion in the number of bilateral trade agreements in the years after the establishment of the WTO in 1995 (an ironic outcome in that the creation of the WTO had been regarded as a significant advance in multilateralism). The WTO reported that it had been notified of nearly 300 new PTAs between 1995 and 2008; bilateral agreements accounted for more than three quarters of all PTAs notified and in force. Of particular significance in this context was the abandonment by East Asian states of their previous commitment to nondiscrimination. Whereas only one preferential trade agreement was in place in the region in the mid-1990s (the ASEAN Free Trade Agreement), a decade later more than 80 PTAs either were being implemented or were under negotiation.

Several factors contributed to this growth in bilateralism: the slow progress made in global trade negotiations and perceptions that bringing the Doha Round of WTO negotiations to a successful conclusion would be difficult; the growth in regionalism in other parts of the world (especially the enlargement and deepening of cooperation within the EU), which caused a "domino" effect as other countries sought not to be excluded from the new preferential agreements; and the financial crisis in East Asia in 1997–1998 (which led governments to reconsider prevailing orthodoxies on trade policy).

Proponents of the new bilateralism believe that it offers several advantages over regional and multilateral negotiations: In principle, bilateral agreements should be simpler to negotiate given the small number of parties involved; they also afford an opportunity to negotiate "deeper" cooperation than in the WTO and in doing so can construct a model that provides states with an incentive to emulate its provisions at the global level. Critics, on the other hand, argue that such arrangements are inherently discriminatory, cause distortions because partial and discriminatory liberalization can advantage actors who are not the most efficient producers, politicize trade and bring power considerations to the fore (see the "Unequal" treaties of the 19th century or those between Nazi Germany and its Central European neighbors), and divert attention and scarce resources from potentially more beneficial negotiations at the global level.

John Ravenhill
Australian National University
Canberra, Australia

See also Diplomacy; International Trade; Multilateralism; Protectionism; Trade Liberalization; World Trade Organization (WTO)

Further Readings

Barton, J. H., Goldstein, J. L., Josling, T. E., & Steinberg, R. H. (2006). *The evolution of the trade regime: Politics, law, and economics of the GATT and the WTO*. Princeton, NJ: Princeton University Press.

Jackson, J. H. (1997). *The world trading system: Law and policy of international economic relations.* Cambridge: MIT Press.

Ravenhill, J. (2002). A three bloc world? The new East Asian regionalism. *International Relations of the Asia-Pacific*, *2*(2), 167–195.

Winham, G. R. (1992). *The evolution of international trade agreements*. Toronto, ON, Canada: University of Toronto Press.

Biology and Politics

For most of the past century, the so-called standard social science model (SSSM) has been the overwhelmingly dominant paradigm in American political science. As we know, the SSSM holds three tenets to be critical:

1. that humans have no innate behavioral tendencies;
2. that, consequently, human nature and human behavior are solely the products of learning and socialization; and
3. that, consequently, human nature and behavior are essentially malleable.

The launching of the "biology and politics" movement in the mid-1960s by a handful of political scientists was the result of two intellectual convictions they held in common. One was an acute dissatisfaction with the state of the discipline in general and with the hegemonic SSSM in particular; the other was a growing appreciation of the significance for political science of recent advances in the biological sciences—especially in neo-Darwinian evolutionary theory, genetics, psychopharmacology (the demonstrated ability of some drugs to change behavior), primatology, and ethology. The first conviction led to a profound difference with their mainstream colleagues in their understanding of the factors that shape human nature and human political behavior. The second produced equally sharp differences with regard to the proper scope and method of political science research and the third in their respective understanding of what constitutes sound public policy. These three, then, constitute the major points of disagreement between proponents of what is commonly called *biopolitics* and the loyal partisans of the SSSM.

The Wellsprings of Political Behavior

As noted above, the SSSM tenet that culture alone shapes behavior is flatly rejected by those who believe in a "more biologically oriented" political science. Yes, culture is important, they agree, but so are the genetically transmitted behavioral tendencies that our species, as social primates, has evolved over literally millions of years. To ignore or even to slight these is to turn a blind eye to forces that often powerfully influence how we act socially and politically. If the discipline ever hopes to understand and explain political behavior, they insist, both nature and nurture must be given serious attention. That calls, in turn, for some major changes in both the scope of what we study and the methods employed in our research.

Research

Expanding the Scope

The Importance of Genetic Influences

Mainstream political science, to be sure, rejects this idea entirely. From an evolutionary stance, like it or not, there are a number of genetic predispositions that we share with other social primate species. Among the politically most important of these are a proclivity for hierarchical social and political structures characterized by dominance and submission, marked differences of status resulting in unequal access to the good things of life, aggressive behavior (especially among males), xenophobia, and nepotistic favoritism.

Relevance of Primatological Data

According to the SSSM, since behavior is shaped entirely by culture, knowledge of other species, no matter how akin to ours, serves no useful purpose. From a biopolitical viewpoint, however, the study of closely related species—and especially the social primates—can yield valuable clues as to the possible sources of our own behavior. This being the case, careful attention should be given to the other apes and, especially, to the species with which we share some 98% of our genetic material, the chimpanzees.

Better Methods

The study of human behavior from a neo-Darwinian perspective emerged from ethology, a biological discipline guided by the dictum that if we hope to understand how and why an organism acts as it does we must study its actual behavior in its natural setting. This means that while there are many similarities in research methodology between the two approaches there are also some key differences. We will mention only two here:

Survey Research

The strengths and shortcomings of survey research as a means of ascertaining attitudes, possible future behavior, or reporting past actions have already generated an abundant literature, so we will be quite brief. As John Wahlke (1979) argued in his American Political Science Association (APSA) presidential address, survey data, which are necessarily based on self-report and with responses often possibly colored by the framing of the questions, are too often unreliable guides to the respondents' actual political behavior, political beliefs, or the intensity of those beliefs. Accordingly, biopolitics, with its ethologically derived focus on actual behavior, sees surveys, by and large, as a research instrument of reluctant resort.

Experiments

Mainstream political science and biopolitics alike view experiments as a valid means of inquiry. The latter insists, however, that the experiments mirror, to the greatest degree possible, the challenges, problems, and environments that its subject(s) would encounter under real-life conditions. Accordingly, its practitioners are skeptical about the validity of experiments where subjects (often college students, hardly representative of the general population) are placed in a patently artificial setting and asked to perform tasks that lack meaningful consequences.

The Formulation of Public Policy

Not surprisingly, the disagreements on the wellsprings of human behavior, on the need to study closely related species, and on key aspects of research methodology result in markedly dissimilar conceptions of how to best shape public policy. Two of the most important of these are the view that past behavior can guide future behavior and that it is able to be converted into policies.

Past Behavior as a Useful Clue to Future Action

Evolutionary theory assumes that, barring mutation and/or some profound environmental change, the behaviors that a species has evolved and exhibited in the past are likely to remain essentially the same. Thus, as those in biopolitics see the world, when efforts to alter or even prohibit "undesirable" social and political behaviors have consistently failed in the past, similar policies are unlikely to be any more successful in the future. In short, the leopard does not readily change its spots. On the contrary, by believing in the essential malleability of human nature, the SSSM is much more prone to discount past failures. Maybe by better manipulating its environment, we can induce, or even better, compel, the leopard to change its coloration.

Trade-Offs Versus "Solutions"

Evolutionary change almost always entails a trade-off—that is, a price. A massive protective shell or carapace comes at the cost of lessened mobility and greater size requires greater caloric demands; as we have seen with our own species, larger brains carry the cost of more painful and more hazardous childbirth. Given their neo-Darwinian set of mind, those in biopolitics tend to approach public policy problems less in terms of far-reaching "solutions" and more in terms of what is to be gained—and what might possibly be the negative political or social consequences. These are matters almost totally ignored, to take some obvious instances, by the ill-fated Eighteenth Amendment, the futile "War on Drugs," and the predictably futile efforts, literally over the centuries, to prohibit prostitution.

There could be no more convincing an example of the "big-solution" approach than the decision to forcibly export "democracy" to Iraq. Given Homo sapiens' innate inclination to hierarchical and authoritarian social and political structures, an inclination evidenced both by the rarity of democracies historically and their relative scarcity today in a so-called Age of Democracy, this mode of governance requires a complex of "special enabling conditions" for its emergence and survival. But even to the most untutored eye, very, very few of these essential, enabling conditions were—or even now are—present in that unfortunate nation.

Slow Start, Brightening Prospects: Long-Term Implications for Political Science

The attempt of the biology and politics movement to reshape the discipline quickly received formal recognition from the International Political Science Association in 1973 and, a bit less readily, from the American Political Science Association in 1981. The task of achieving a distinct "field" identity was basically completed with the establishment of a professional organization, the Association for Politics and the Life Sciences in 1980; the establishment of a peer-reviewed journal, *Politics and the Life Sciences*, in 1982; and the publication, several years later, of a 90-page Master Bibliography listing more than 1,000 "biopolitical" papers, articles, and books.

Nonetheless, their determined evangelical efforts notwithstanding, until quite recently the proponents of a "more biologically oriented" political science constituted only a minuscule fraction of the discipline, with the Association for Politics and the Life Sciences having, at most, only 300 or 400 members in contrast with APSA's some 101,000 members. Understandably, the great majority of the profession was unwilling to consider, let alone embrace, a conceptual framework at loggerheads with the orthodoxy in which they had been trained and had labored. Although a biopolitical course or two eventually made its way into the graduate curriculum at a dozen or so schools, only one department (at Northern Illinois University) initiated a formal doctoral program in the field; very few of the professional journals published biopolitical articles; with rare exceptions, biopolitical books somehow escaped the attention of the journals' book review editors; and the one or two biopolitical panels at APSA's annual meeting were fortunate to draw more than a dozen or so attendees.

Over the past 3 or 4 years, however, there has been a definite sea change. The professional journals have become much more receptive to biopolitical topics; both the number of biopolitical panels and panel attendance at the APSA meetings have notably increased; and a biopolitically oriented Conference on Experimental Political Science held its third annual meeting in 2010.

There are several reasons for this development. The introduction of functional magnetic resonance imaging (MRI) technology has opened a new way to study how the brain reacts when experimental subjects are presented with political questions or stimuli—and this research has received a surprising, if regrettably uncritical, coverage in the popular media (some relevant works on the neuropolitical perspective can be found in Rose McDermott, 2007, and Darren Schreiber, 2007). Another handful of biopolitical papers have reignited interest in the debate about the respective roles of nature and nurture in shaping political attitudes (in this instance, of identical twins)—and these, too, have received flattering media attention.

In addition, other influences have also been operative. Over the past couple of decades, biologists have done a superb job of popularizing evolutionary theory and the understanding of, and interest in, the manner in which genes and culture interact to shape our behavior. At the same time, an evolutionary approach has made considerable headway in several of the other behavioral sciences—anthropology, sociology, psychology, law, and what is now called "behavioral" economics. Political science has a well-established history of borrowing concepts and methods from our sister disciplines, and, among these, psychology and economics have long been the most prestigious and influential. And we should surely mention one more possible influence. According to Thomas Kuhn (1970), new paradigms often triumph not simply because of their superior scientific merit but also because of the implacable toll taken by rigor mortis on the avatars of the old order.

Whatever the relative weight of these several factors, their combined impact will almost surely profoundly modify our discipline's understanding of the forces shaping political behavior, the manner in which it conducts its research, and the intellectual framework within which it formulates its public policy recommendations. In the future, advocates of these views may come to represent a sizable plurality of the profession rather than the current small minority.

Albert Somit
Southern Illinois University
Carbondale, Illinois, United States

See also Behavioralism; Democracy, Theories of; Policy Analysis; Political Psychology

Further Readings

Dawkins, R. (2006). *The selfish gene* (3rd ed.). New York: Oxford University Press.
Degler, C. (1991). *In search of human nature.* New York: Oxford University Press.
Hibbing, J., Alford, J. R., & Funk, C. L. (2005). Are political orientations genetically transmitted? *American Political Science Review, 99,* 153–166.
Kuhn, T. (1970). *The structure of scientific revolutions.* Chicago: University of Chicago Press.
McDermott, R. (2007). Cognitive neurosciences and politics: Next steps. In W. R. Neuman, G. E. Marcus, A. N. Crigler, & M. MacKuen (Eds.), *The affect effect: Dynamics of emotion in political thinking and behavior* (pp. 375–398). Chicago: University of Chicago Press.
Schreiber, D. (2007). Political cognition as social cognition: Are we all political sophisticates? In W. R. Neuman, G. E. Marcus, A. N. Crigler, & M. MacKuen (Eds.), *The affect effect.* Chicago: University of Chicago Press.
Somit, A., & Peterson, S. A. (1990). *Biopolitics and mainstream political science: A master bibliography.* DeKalb, IL: Program for Biosocial Research.
Somit, A., & Peterson, S. A. (1997). *Darwinism, dominance, and democracy.* Westport, CT: Praeger.
Somit, A., & Peterson, S. A. (2005). *The failure of democratic nation-building.* New York: Palgrave Macmillan.
Wahlke, J. (1979). Pre-behavioralism in political science. *American Political Science Review, 73,* 9–32.

Bipolarity and Multipolarity

Bipolarity and multipolarity are theoretical concepts that refer to the distribution of relative power capabilities among the major powers in the global and/or regional international system. Bipolarity exists when there are two dominant powers—commonly referred to as "superpowers"—in the system whose power capabilities are considerably greater than those of other major powers. The prime example of bipolarity is the U.S.–Soviet competition during the Cold War. Multipolarity exists when there are three or more great powers in the system. Historically, multipolarity has been the dominant power configuration in the modern European states' system.

Polarity: Theoretical Approaches

The significance of the distribution of relative power capabilities in the international system is widely contested. For realists, it is the single most important determinant of international politics. Liberals also accept that the distribution of power is an important factor shaping international politics, although they attach greater significance to domestic-level factors (such as regime type) or the existence of international regimes and institutions as determinants of international outcomes. Similarly, the English School has traditionally recognized the importance of the balance of power and the management role of great powers, even though this approach stresses the importance of "international society" as a factor that ameliorates the operation of power politics.

Critical theorists and postpositivists, on the other hand, question the significance of polarity as a theoretical tool for understanding international politics. These, more radical approaches challenge the very concept of "power" as used by realists and other rationalist theories; power, they argue, cannot be reduced to resources, geography, and other material factors. Like constructivists, they tend to emphasize the significance of normative and ideational factors and stress the role of nonstate actors. Consequently, they argue that concepts of bipolarity and multipolarity are irredeemably freighted with realist assumptions and that focusing on polarity obscures the "real" operation of power in the international system that serves the interests of privileged elites.

For realists, however, the distribution of power in the international system is the key factor shaping any regional or global conjuncture. International political systems are inherently unequal, diverse, and pluralistic: They consist of a variety of different actors with wildly differing power capabilities, with a relatively small number of more powerful states. Realists argue that this handful of powerful states—generally known as great powers—exert a disproportionate influence on international affairs and that relations between them provide the key determinants of the structure and dynamics of the international system.

Classical Realism

The classical realist tradition of Hans Morgenthau, E. H. Carr, and Reinhold Niebuhr, as well as English School theorists such as Hedley Bull and Martin Wight, were centrally concerned with the balance of power. They explored the relations between the great powers from a perspective informed by the study of history and philosophy, and their analyses of multipolarity and bipolarity drew heavily on the experience of the modern European states' system. Much of their work explored the implications of multipolarity and bipolarity for diplomacy and statecraft. Classical realists tended to regard multipolar systems as more stable than bipolar systems, because their greater uncertainty and unpredictability encouraged foreign policy decision makers toward caution and circumspection.

Waltz's Theory of International Politics

With the publication of Kenneth Waltz's *Theory of International Politics* in 1979, the concepts of multipolarity and bipolarity acquired a new theoretical context and analytical significance. Waltz sought to introduce a far-reaching "Copernican revolution" into the study of international relations by demonstrating how much of states' actions and interactions, and how much of the outcomes their actions and interactions produced, could be explained by forces operating at the level of the system rather than at the level of the units. His neorealism (also known as structural realism) was a parsimonious theory that focused almost exclusively on the distribution of relative power capabilities between states, specifically the great powers. Waltz argued that the determining feature of any international system, and the basis for any theoretical analysis of international politics, must be about how power is distributed between the major units. If there were three or more great powers, the system was characterized by multipolarity; if there were only two great powers—or superpowers—the system was bipolar. This distribution of relative power capabilities was thus crucial in determining international political outcomes and shaping the structural conditions for war or peace.

Waltz's *Theory of International Politics* was as controversial as it was influential. In contrast to many of his classical realist predecessors, Waltz argued that multipolar systems were more unstable and war prone than bipolar systems because in the former, the great powers had a number of potential conflict dyads and defection from an alliance could have a destabilizing impact on the balance of power. In multipolar systems, he argued, dangers were diffused, responsibilities unclear, and definitions of vital interests easily obscured. Moreover, in multipolarity, it was not always clear who was a danger to whom, which complicated the process of alliance building and the establishment of a stable balance of power. In bipolar systems, in contrast, there were only two superpowers and, consequently, the question of who was a danger to whom was never in doubt. This, he argued, was the first big difference between bipolarity and multipolarity. Moreover, power balancing between them depended primarily on "internal balancing" rather than on "external balancing" (i.e., on the generation and more effective use of their own capabilities and resources rather than on alliance building, which relies on the capabilities of others). Thus, the defection of allies was of less significance than in multipolar systems.

Waltz, however, also argued that the "default" position of international systems, whether bipolar or multipolar, was the balance of power. This, he argued, constituted the key insight of his theory of international politics. From the theory, he argued, one could predict that states will engage in balancing behavior, whether or not they consciously seek a balance of power. The expectation, he argued, is not that a balance, once achieved, will be maintained but that a balance, once disrupted, will be restored in one way or another. Waltz's argument that international political systems tend toward balance, and that states were primarily concerned with maximizing their security rather than their power, gave his brand of structural realism a relatively benign and status quo flavor, leading some of his realist critics to dub his theory "defensive realism."

Neorealist Critiques of Waltz

First and foremost among Waltz's Neorealist critics was John Mearsheimer. He argued that precisely because great powers were security maximizers, they were also power maximizers. In a competitive, self-help system, Mearsheimer argued, the most effective way for great powers to ensure their security was by becoming the most powerful state in their neighborhood and establishing their regional hegemony. Hence, great powers, whether they liked it or not, were faced by a systemic imperative to maximize their power. This emphasis on power maximization is at the heart of Mearsheimer's theory of "offensive realism," which presents a much more conflict-prone and pessimistic view of the "tragedy" of great power politics.

Mearsheimer also introduced an important distinction between different forms of multipolarity. He argued that where power capabilities are roughly equivalent between the great powers, a situation of "balanced multipolarity" existed in which no one great power could make a bid for regional hegemony. In this context, great power cooperation was possible, as occurred in the "Concert of Europe" following the end of the Napoleonic Wars. However, in the context of a multipolar system in which one great power was significantly stronger than the others, "unbalanced multipolarity" existed. Unbalanced multipolar systems were inherently war prone, because the strongest great power would face a systemic imperative to maximize its power by making a bid for regional hegemony. For Mearsheimer, his theory of "offensive realism" provided an explanation for the major great power conflicts, such as the Napoleonic Wars and World Wars I and II.

The Contemporary Debate

With the end of Cold War bipolarity, a wide-ranging debate has taken place on the nature and significance of the distribution of relative power capabilities in the international system. One thread of this debate focuses on the concept of "unipolarity." Some realists have argued that the United States, as the sole remaining superpower, is so overwhelmingly powerful compared with other great powers that the system is best described as "unipolar." Among American realists, much of the debate has focused on whether this "unipolar" world order will be relatively short lived or long lasting (i.e., whether it constitutes a unipolar "moment" or a unipolar "era"). Waltz, on the other hand, has argued that after a relatively brief interregnum in which former great powers "relearn" their great power roles and new great powers emerge, the system will return to a multipolar dynamic. In this context, the emergence of the "BRIC" countries (Brazil, Russia, India, and China) has been taken by some as heralding the emergence of a new multipolar international order.

Against these debates among realists on the nature of contemporary polarity, more radical approaches have taken issue with the realist understanding of the nature of power and power politics. Power is an essentially contested concept in political science, and some have argued that power cannot be understood in terms of material capabilities. Rather, they argue that nonmaterial factors such as reputation, trust, and the attractiveness of states' economic, political, social, and cultural attributes constitute a new form of power, which they describe as "soft power." Some critical theorists even speak of "normative power" (i.e., the power of norms and values), which, they suggest, can have a system-transforming impact over the long term.

These more radical understandings of power are linked to arguments about the changed nature of international politics. Some argue that economic interdependence and the emergence of a globally integrated international order have weakened the significance of the distribution of hard power capabilities and therefore made long-established notions of multipolarity and bipolarity largely redundant. Globalization, some now argue, has blurred the distinction between domestic and international politics and weakened the significance of sovereignty. In this postmodern system, borders are increasingly porous, power is diffused, and security can be achieved through governance and interdependence rather than through the balance of power. Concepts of bipolarity and multipolarity thus continued to be employed in international relations, although their significance is deeply contested.

Adrian Hyde-Price
University of Bath
Bath, United Kingdom

See also Balance of Power; Cold War; Globalization; Power; Superpower

Further Readings

Hyde-Price, A. (2007). *European security in the twenty-first century: The challenge of multipolarity.* Abingdon, UK: Routledge.

Little, R. (2007). *The balance of power in international relations: Metaphors, myths and models.* Cambridge, UK: Cambridge University Press.
Mearsheimer, J. (2001). *The tragedy of great power politics.* New York: W. W. Norton.
Mogenthau, H. (1989). *Politics among nations: The struggle for power and peace.* New York: McGraw-Hill.
Waltz, K. (1979). *Theory of international politics.* New York: McGraw-Hill.

BOOLEAN ALGEBRA

Boolean algebra is the subfield of mathematical logic devoted to operations and functions of logical variables. After a brief historical introduction, a broad idea of its main concepts, ideas, rules, and applications is given below.

Origins

In his work about the laws of thought, the English mathematician George Boole (1815–1864) showed that it is possible to represent logical thought processes using a system of binary variables: 0-1, TRUE-FALSE, and YES-NO.

As early as 3000 BCE, the Chinese emperor Fu Hi used, within a magical symbol, a set of eight trigrams, in which three binary variables yield a written representation of integers from 0 to 7 (Figure 1). Gottfried W. Leibniz (1646–1716) took inspiration from it when he laid down the basic rules of binary arithmetic (*Explanation of Binary Arithmetic*, 1703). He realized that he had found a way to translate logic into calculations and laid plans for the first four-operation computing engine. However, it needed Boole's work for the idea to mature, giving birth to a genuine theory—Boole's algebra—with its calculation rules, its properties, and its theorems.

Sets, (Classical) Logic, and Boolean Algebra

A specificity of this new theory is that it may be explained under three completely equivalent forms, notwithstanding their apparent differences, using set theory, elementary logic, or Boolean algebra. In a way, we have three different descriptions of the same theory, in three different languages, but allowing perfect translation between them.

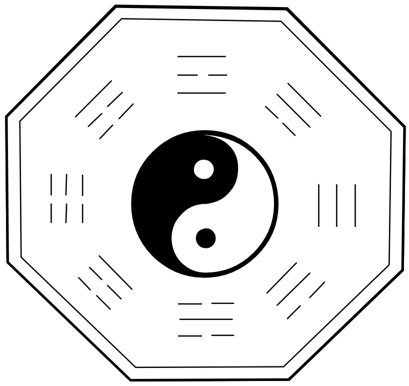


Figure 1 Fu Hi Trigrams (From Ancient China)

Sets

"Naive set theory" uses a set of well-defined elements—its universe—and considers a series of relations and operations between those elements and subsets of this universe. For example, if the universe is the set of all national flags at some specific moment in time, one can easily imagine all kinds of subsets: flags using red, black, or both; those featuring a cross, a circle, bars, stripes; and so forth. Using red is a binary variable, with value YES (or 1) or NO (or 0) for each flag.

Although not absolutely needed for conceptual development, a visual representation using Venn diagrams can show this relationship (see Figure 2). This figure shows the flag Universe and three of its subsets, namely flags using red (R), those using black (B), and those using yellow (Y). It immediately can be seen in the diagram that each subset automatically determines its complement (flags without black, without red, or without yellow). The Italian flag (i), for example, belongs to subset R but to neither B nor Y; this can be noted as $i \in$ R, $i \notin$ B, and $i \notin$ Y.

It is also possible to consider operations on subsets. The *intersection* ($\cap$) of **R** and **B** is the subset of all flags using both red AND black; the *union* ($\cup$) of **R** and **B** is the subset of all flags using either red OR black—or both; this is the inclusive variant of OR. One can see that the French flag (f) satisfies $f \in \mathbf{R} \cup \mathbf{B}$ and $f \notin \mathbf{R} \cap \mathbf{B}$. Finally, some subset may be totally *included* ($\subset$) within another: The set of

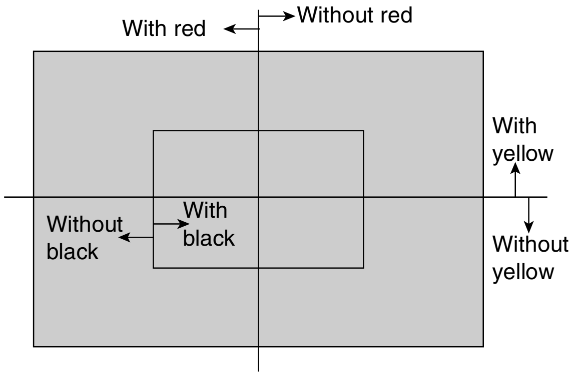


Figure 2 Venn Diagram With Three Subsets of Flags

all flags using both red and black is included in the set of all flags that merely use red. This is noted as **R ∩ B ⊂ R**.

Logic

Classical logic is able to translate any statement of the elementary set theory; for example, **R ∩ B ⊂ R** translates into an implication (⇒): If a flag uses both red and black, then it uses red. More generally, any statement of the form A ⊂ B becomes A ⇒ B (if A is true, then so is B). Three logical operators are defined at the beginning: logical conjunction, "AND"; logical disjunction, "OR"; and negation (or denial), "NOT."

Boolean Algebra

Boolean algebra is based on the following conventions: Each variable is binary; it has only two possible values; and those will be noted 0 (corresponding to the empty set and logical FALSE) and 1 (corresponding to the universe and logical TRUE).

Boole then defined three operations: Boolean product (·), Boolean sum (+), and complement. Boolean product corresponds to logical AND and a set-theoretical intersection; Boolean sum corresponds to OR and set-theoretical union; and complement corresponds to set-theoretical complement and logical NOT.

To determine how operators make this algebra work, the rules of the game—that is, the axioms—are needed. They are the following:

Law of excluded middle: $A + \bar{A} = 1$

Noncontradiction principle: $A \cdot \bar{A} = 0$

Boundedness conditions: $A \cdot 1 = A$; $A + 0 = A$

Twofold commutativity: $A + B = B + A$; $A \cdot B = B \cdot A$

Twofold associativity: $(A + B) + C = A + (B + C)$; $(A \cdot B) \cdot C = A \cdot (B \cdot C)$

Reciprocal distributivity of "+" and "·": $(A + B) \cdot C = (A \cdot C) + (B \cdot C)$; $(A \cdot B) + C = (A + C) \cdot (B + C)$; this last identity is less intuitive than others, since it does not hold in ordinary (real-number) arithmetic.

From there, several properties may be deduced, among which are the following:

$\bar{0} = 1$ *and* $\bar{1} = 0$

Idempotence: $A \cdot A = A$; $A + A = A$

Absorbing: $A + 1 = 1$ and $A \cdot 0 = 0$

Successive applications of these rules make possible the simplification of complex Boolean expressions, in particular by using the following results:

Absorption theorems: $A + (A \cdot B) = A$; $A \cdot (A + B) = A$

Redundancy theorems: $(A \cdot B) + (\bar{A} \cdot C) = (A \cdot B)$ $(\bar{A} \cdot C)(B \cdot C)$

De Morgan's laws: $\overline{A + B} = \bar{A} \cdot \bar{B}$; $\overline{A \cdot B} = \bar{A} + \bar{B}$

Boolean Functions and Truth Tables

Any Boolean function—that is, any combination of Boolean variables using Boolean operators—may be analyzed using its truth table. The truth table for a function of n variables is always made of $2n$ lines and $n + 1$ columns; each line corresponds to one combination of values of the variables (0 or 1). For example, the truth table for function $F(A, B) = A \cdot \bar{B} + \bar{A} \cdot B$. B is as follows (Table 1):

Table 1 Truth Table for $F(A, B)$

A	*B*	*F(A,B)*
0	0	0
0	1	1
1	0	1
1	1	0

All truth tables of two variables will have the same first two columns as this one; the third column completely describes the function. This last column is made up of four binary values; therefore, there are 24 = 16 different functions of two variables. Using set-theoretical representation, function $F(A, B) = A \cdot \bar{B} + \bar{A} \cdot B$. B is shown in the shaded part of Figure 3.

Using two variables, 16 different shaded zones may be defined, made up of 0 to 4 elementary zones (the quadrants of the diagram) (Table 2).

This number increases rapidly with the number n of variables.

This exponential increase explains the need for using algorithms when analyzing Boolean functions of more than three variables (see Figure 3).

Each compound zone of a Venn diagram also corresponds to a Boolean function. However, there is more than one formula for a given function.

It is therefore interesting to find, among those logically equivalent formulae, which one is the simplest. Simplicity is not a well-defined notion, but an intuitively reasonable approximation will be to choose the shortest formula.

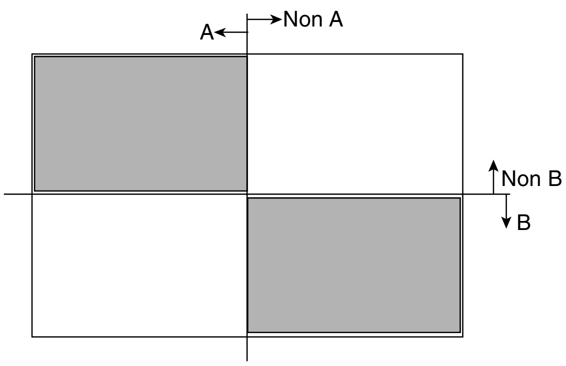


Figure 3 Venn Diagram With Two Subsets

Table 2 Number of Zones

Number of Variables	*Elementary Zones*	*Compound Zones*
1	2	$4 = 2^2$
2	4	$16 = 2^4$
3	$8 = 2^3$	$256 = 2^8$
n	2^n	$2^{(2n)}$

Reducing a Function's Algebraic Expression

The longest formula for a given function (its canonical disjunctive form) is very easy to write down: It is merely made up of a union of elementary zones, therefore of a sum of products (monomials, minterms) using all variables. For example, to the following truth table (see Table 3) corresponds the long formula $F(A, B, C) = \bar{A} \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C + A \cdot \bar{B} \cdot C + A \cdot B \cdot C$.

This formula lacks brevity. Using a succession of Boolean algebra rules allows us to simplify it until we get the shortest possible formula:

$$F(A, B, C) = A \cdot B \cdot C + \bar{A} \cdot B \cdot C + A \cdot \bar{B} \cdot C + \bar{A} \cdot \bar{B} \cdot C$$

$$F(A, B, C) = 1 \cdot B \cdot C + 1 \cdot \bar{B} \cdot C$$

$$F(A, B, C) = B \cdot C + (\bar{B} \cdot C) = (B + \bar{B}) \cdot C = 1 \cdot C = C.$$

This search for the shortest form of the function is the object of the simplification algorithms mentioned above. Karmaugh's algorithm, a heuristic method for simplifying that can be worked by hand, with acceptable results, for $n \leq 4$, may give a good idea of how such algorithms work; other algorithms need a computer to perform, such as the

Table 3 Truth Table for $F(A, B, C)$

A	*B*	*C*	*F(A,B,C)*
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

Quine-McCluskey algorithm, used in qualitative comparative analysis software.

Gisèle De Meur
Université Libre de Bruxelles
Brussels, Belgium

See also Comparative Methods; Configurational Comparative Methods; Qualitative Comparative Analysis

Further Readings

De Meur, G., & Rihoux, B. (2002). *L'analyse quali-quantitative comparée* [Comparative quali-quantitative analysis]. Louvain-La-Neuve, Belgium: Academia Bruylant.

Ragin, C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.

Rihoux, B., & Ragin, C. C. (Eds.). (2009). *Configurational comparative methods: Qualitative comparative analysis and related techniques.* Thousand Oaks, CA: Sage.

Bootstrap

See Robust Statistics

Breakdown of Political Systems

The term *breakdown of political systems* connotes the collapse of any rule-based institutional network engaged in the allocation of scarce resources. Though this rubric could technically embrace networks ranging from empires to smaller, nonstate entities such as clans, the vast majority of the literature on the breakdown of political systems focuses on the collapse of either regimes or states. The disintegration of both regimes and states has been a constant throughout history, and thus, the literature on the breakdown of political systems is vast. This overview merely introduces this literature through a discussion of its basic terms, its origins, and some of its principal arguments. The discussion is confined to four types of political system breakdown: the breakdown of dictatorship, the breakdown of democracy, and the breakdown of states, leading either to new, functioning states or to failed states.

The Basic Notions and Its Origins

Drawing on Robert Fishman's distinctions, we can define a *regime* as a country-based network of formal and informal organizations at the center of political decision making that determines who has access to political power and how those who are in power deal with those who are not. Democracies and dictatorships are the dominant regime types in our literature.

A *state* is the larger institutional context in which a regime exercises its power. A state is a more permanent structure of domination and coordination in part because, unlike a regime, a functioning state always maintains a coercive apparatus and both the means to administer a society and to extract resources from it. Though there are many state types in our literature, the distinction most relevant to a discussion of system breakdowns is the one between functioning states and failed states.

The current literature on the breakdown of modern political systems has its roots in ancient Athens. Plato attributed political system change (and difference) to "the dispositions of men" and argued that regime types were forged from "the human natures" of their citizens. Later classic commentaries linked the breakdown of oligarchy to societal crises related to taxation, to the successful uprising against the Spartan occupation that followed conflict between Athenian elites, and to structural changes related to the Persian Wars. Aristotle linked political system longevity to levels of economic inequality. Arguments relating the breakdown of political systems to cultural dispositions, fiscal crises, elite divisions, war, and economic inequality remain fundamental to political science literature today.

The Ancients' attempts to derive general, explanatory arguments from specific, contemporary political events remain fundamental to current literature as well. This is one of the reasons why scholarship on the breakdown of political systems has so often come in clusters, following waves of dramatic political system change. Four clusters of system breakdown stand out as catalysts.

The Main Explanations

The first set of debates about regime collapse in the 20th century focused on the breakdown of democracy in interwar Europe. Proponents of psychosocial arguments asserted that the cause lay in the nexus of personality and politics: Individuals who were profoundly insecure found the uncertainty of democracy intolerable and sought comfort in authoritarian leaders and authoritarian regimes. Though this argument was rooted in social psychology, it resonated with scholars of politics as well. Hannah Arendt attributed the breakdown of interwar democracy to the "negative solidarity" engendered by a mass of "furious individuals" united by "apprehension." Seymour Martin Lipset attributed "working-class authoritarianism" to economic insecurity and authoritarian family patterns.

While some analysts concentrated on citizen attitudes as a root cause of democratic breakdown, other scholars concentrated on classes instead. Barrington Moore attributed the breakdown of interwar democracy to an alliance between landed elites and sectors of the bourgeoisie who desired a level of labor repression that democracy would not allow. Gregory Luebbert attributed the breakdown to alliances between the family peasantry and the urban middle classes, arguing that their interactions with socialist parties were decisive.

The focus on class actors was eventually challenged by approaches that focused on political institutions and political (rather than class) elites. Juan Linz and Alfred Stepan took the lead in fostering this approach. Though Linz ultimately blamed the breakdown of democratic political systems on "citizens of weak commitment," he attributed citizen defections to a broad range of nonstructural factors, including ineffective leadership and problematic institutional design. In so doing, he argued that system breakdown might have been avoided with better judgment and better institutions, including single parliamentary executives, constructive votes of no confidence, and laws preventing the proliferation of small parties.

The breakdown of democracy occurring during the Cold War was the second major stimulus for explanatory theory and Linz intended his argument to apply to these cases as well. Gabriel Almond and Sidney Verba also sought to offer a general theory of democratic breakdown but promoted arguments related to political culture instead. They proposed that democratic stability could only be ensured where a "civic culture" based on moderate attitudes, interpersonal trust, and muted interest in political participation existed. In a similar vein, Harry Eckstein proposed that political systems would be viable only if their authority patterns were congruent with authority patterns in other social organizations.

Suspecting that the roots of citizen attitudes and behaviors lay in economic structures, Lipset proposed that a democratic system breakdown would be less likely as economic development progressed, arguing that the prospects for democratic system stability improved as wealth, industrialization, urbanization, and education expanded. A large middle class would mean decreasing class differences, more moderation, and better odds for democratic system stability.

Focusing on the frailty of democratic political systems in South American countries with comparatively large middle classes, Guillermo O'Donnell challenged Lipset and argued that economic development would not guarantee democratic stability if it took place in the context of economic dependency. Democracy could not be maintained in the face of contradictory pressures from transnational capital on the one hand and a popular sector empowered by democratic freedoms on the other.

Debates about the cause of democratic system breakdown were still unresolved when authoritarian regimes began to break down in the 1970s. The theories stimulated by this third set of events were equally diverse. Philippe Schmitter and Guillermo O'Donnell offered one of the first attempts to make sense of the change. They proposed that the breakdown of dictatorship is always the consequence of divisions among regime elites and that "elite dispositions, calculations, and pacts" determine whether system breakdown is possible. Beyond this, they insisted that the breakdown of authoritarianism constituted an example of "underdetermined" social change where class, sectoral, or institutional analysis had little predictive utility.

As cases of authoritarian breakdown began to accumulate, scholars argued that the causes of regime breakdown varied across countries. Samuel Huntington drew attention to three scenarios for breakdown: one in which regime elites played a major role, another in which the breakdown of dictatorship resulted from joint action by regimes and oppositions, and a third in which oppositions took the lead. This last scenario dominated in Africa and Eastern Europe.

Barbara Geddes pointed out that different sorts of authoritarian systems broke down for different reasons and at different rates. Through quantitative work, she found that military dictatorships were the most likely to break down, that personalist regimes were more resilient, and that single-party regimes were the most resilient of all. Military regimes were susceptible whenever the interests of the military as an institution were at risk. Personalist regimes were susceptible when an equilibrium based on coercion and spoils was disturbed by the death of the leader or by economic crisis. Single-party regimes were the least susceptible to breakdown because their leaders typically had no alternative source of power and thus had stronger incentives to face crises and resolve disputes internally.

The idea that economic crises or economic factors of some sort are the root cause of the breakdown of dictatorship is widespread but still contested. Studying the period between 1950 and 1990, Adam Przeworski and Fernando Limongi found that the relationship between economic development and the breakdown of dictatorship was bell shaped. Though the likelihood of regime breakdown rose with development in middle-income dictatorships, dictatorships in countries with either very low or very high incomes tended to be stable despite growth. Studying the period from 1800 to 1990, Carles Boix and Susan Stokes drew a different conclusion. They found that economic development always *increased* the probability that dictatorships would become democracies, though the impact of growth on the likelihood of regime change decreased at higher levels of development and after World War II.

Debates about the roots of dictatorial political system breakdown were enriched by the wave of regime changes associated with the collapse of the Soviet Union in 1989. This fourth stimulus for explanatory theory brought arguments with an international focus to the fore. Scholars such as Marc Beissinger showed how the breakdown of dictatorships was affected by the cross-border diffusion of models for mobilization. Huntington argued that demonstration effects were so powerful that they worked independent of economic and social conditions. Jon Pevehouse illustrated that the breakdown of dictatorship was linked to membership in international organizations dominated by democracies. Dankwart Rustow concluded that the most powerful impetus for the breakdown of dictatorship was the global trend of "intensifying communication and economic integration," which spread the awareness of "democratic lifestyles" and delegitimated the myths that sustained dictatorship.

In addition to invigorating work on the breakdown of regimes, the disintegration of the Soviet Union sparked new interest in theories explaining the breakdown of state systems. Scholars explained why new states emerged from old states with a variety of arguments. Linz and Stepan attributed the change to problems of "stateness": Subnationalist groups that never accepted the boundaries of the state moved to change them as soon as coercion declined. Valerie Bunce offered an institutional explanation, highlighting the fact that all the European states that split apart after 1989 were federal. Carol Skalnik Leff made the point that democratization had a "state-shattering" impact because it shifted power to regions and thereby changed the incentives of regional elites.

The system breakdowns after 1989 also sparked new literature explaining the emergence of "failed" states as places where state actors prey on their citizenry, enjoy no monopoly of force, and fail to control rampant violence. This form of system breakdown is most likely in poor countries, but Robert Bates points out that its roots may lie in poor state resources rather than in poverty per se. Robert Rotberg argues that poverty and institutional weakness contribute to state failure but accords the greatest weight to leadership errors and the drive for personal gain. Jack Goldstone, Robert Bates, Ted Robert Gurr, and Monty Marshall conclude that state failure is most rooted in political factors and that hybrid regimes that are neither democracies nor dictatorships are most susceptible to collapse. We cannot know which of the many explanations for system breakdown given here will prove most convincing in the future, but we can be almost certain that political systems will continue to collapse.

Nancy Bermeo
Nuffield College, University of Oxford
Oxford, United Kingdom

See also Democratic Consolidation; Democratization; Development, Political; Dictatorship; Regime (Comparative Politics); Secession; State Collapse; State Failure

Further Readings

Bermeo, N. (2003). *Ordinary people in extraordinary times: The citizenry and the breakdown of democracy.* Princeton, NJ: Princeton University Press.

Bratton, M., & van de Walle, N. (1997). *Democratic experiments in Africa: Regime transitions in comparative perspective.* Cambridge, UK: Cambridge University Press.

Dawisha, K., & Parrott, B. (1997). *Authoritarianism and democratization in postcommunist societies 1–4.* Cambridge, UK: Cambridge University Press.

Goldstone, J., Bates, R., Gurr, T. R., & Marshall, M. (2005). *A global model for forecasting political instability: State failure task force.* McLean, VA: Science Applications International.

Huntington, S. P. (1991). *The third wave: Democratization in the late twentieth century.* Norman: University of Oklahoma Press.

BUDDHISM

Buddhism is frequently regarded as a quietist religion, but while there may be some justification for this perception, its political presence and impact in many parts of Asia should not be underestimated. Ever since the period of the Indian emperor Aśoka (269–232 BCE), the religion has shown itself to be a distinctive social and political force in various countries, having considerable backing among the relevant populations. Naturally, where Buddhism is a small minority, as in modern Western countries, its influence is small and the quietist image is strong, but in the countries of Asia, where Buddhism is the majority religion, the situation is quite different. Here, both Buddhist monks and lay leaders have been prominent from time to time in elitist and military maneuverings and more recently in revolutionary and electoral situations. Self-consciously Buddhist political activism is not a stranger in India, Burma, Thailand, and Vietnam. Indeed, in most Asian countries, Buddhist leaders have at various times contributed to both support the prevailing establishment and revolt against it.

Characteristics of Buddhism

The underlying reason for the complex interactions between Buddhism and political forces may be found in two characteristic features of the religion. First, we must look back to the story of the historical Buddha's decision to proclaim his teaching in the first place, which is a constitutive narrative both for his own biography and for the emergence of Buddhism as an organized religious system. The central story is that Gautama (in the 5th century BCE), though blessed with many privileges as a young man, passed through a personal heart-searching experience about human suffering, illness, old age, and death. Recognizing the transience of existence, he perceived the cause of suffering to lie in ignorance, desire, and attachment to transient things, while its resolution lay in knowledge and detachment. The understanding of all these relationships is summed up as his "enlightenment" or *bodhi*, which is why he was regarded as a *buddha*, an enlightened one. Had he preferred a quietist, personal resolution of these general questions of human existence, which according to the narrative he considered steadily as an option, the social impact would most likely have been nil. As it was, the very decision to proclaim his system of teaching (in Sanskrit *dharma*, in Pali *dhamma*) led not only to the first ordinations of monks but also to sponsorship by prosperous laity and royalty. This meant that a strong and active symbiosis with a supportive laity grew up from the very beginning, so that the social and political ramifications began to extend themselves in a continuous process.

Second, the concept of the "wheel turner" (in Sanskrit *cakravartin*, in Pali *cakkavattī*) is significant. On one hand, it is part of the mythologization of the figure of the Buddha himself, for he is regarded as having achieved the right to "turn the wheel of dharma." However, the same concept is applied to kingship, for a "wheel-turning monarch" is one who claims the power of extending his rule to the four quarters of the earth, spreading stability in support of a moral society. When a suitable individual is reborn as a "great person," with excellent qualities and at an appropriately high-class of society, he—presumed to be male—has the option of turning the wheel of dharma or of turning the wheel of political power. Ideally, therefore, a powerful and righteous king will be the contemporary and the patron of a living Buddha. This dual, tandem-like concept was later integrated under the concept of the "wonderful union of king and Buddha." In some cases, such as the Tibetan model, this means that the incarnated spiritual leader or Dalai Lama is at the same time the political head of the population, which naturally causes a collision with any modern state such as China, with its secular ideological base. Tensions over this model may be seen to a lesser extent in Mongolia, Bhutan, and Nepal. In a variation of this union, certain kings or princes have been given the title of "Dharma King" (in Sanskrit *dharmarāja*), which has usually implied not only an elitist espousal of Buddhism, but also that they have authority, direct or indirect, over senior appointments in the monastic hierarchy, as frequently seen in premodern Japan.

Evolution of Buddhism in Southeast and East Asia

The expansion of Buddhism into the countries of Southeast and East Asia took place largely with the help of political domination or influence, with the notable exception of China. As a result, these underlying concepts of the Buddha as both political and spiritual head have played a major role. Starting with a response by the Indian Emperor Aśoka to the Sri Lankan king of the time, Devānampiyatissa, the Theravāda monastic order was set up in Sri Lanka. While this order had royal patronage, the tradition of lay support for the order spread to most levels of society. As a result, the postcolonial independence movement looked to Buddhism, still predominating in a religiously and culturally plural society centuries later, to provide an ideological reference point that would be nationally recognizable as Sinhalese. An important modern Sri Lankan Buddhist was Dharmapala (1864–1933), who, assuming the title of Anagarika or "homeless one," did much to further a modern revival of Buddhism in India itself, where it had largely died out. This revival was massively furthered by the conversion of Dr. Babasaheb R. Ambedkar (1891–1956), a major contributor to the constitution of independent India, who led some millions of previous "untouchables" into a casteless state under the banner of a modernist, rationalist, and ethically oriented Buddhism.

Following the earlier mission to Sri Lanka, a transmission of the *Dhamma* to Burma, Thailand, Vietnam, and Laos followed in various stages, with the result that Buddhist traditions played a major role in state formation in all of these countries. There were also establishments in Sumatra and Java that had less long-term influence. While kingship models derived from Hinduism were also influential, with even Brahmans functioning at the Thai court, it was the Buddhist monastic order that has remained a pillar of political influence in mainland Southeast Asia. Thus, regardless of contemporary politicians, the two fundamental foci of authority in Thailand remain to this day the monarchy and the monastic order. A key figure in the formation of Thai national consciousness was King Mongkut (Rāma 4th, 1824–1851) who managed the modernization process of the whole country while preserving its independence from Western colonial powers. Before becoming king, he was a dedicated monk who played a leading role in reforming and modernizing the Buddhist religion in Thailand. The process of modernization was continued by his successor Chulalongkorn (Rāma 5th) and later monarchs. In Burma, Buddhist layman U Nu (1907–1995) played a major role in postcolonial political development under the slogan of "Buddhist socialism," while at the beginning of the 21st century a number of monks have been attempting to recall the country to its rightful path as they see it. In Vietnam, the wars of independence saw cases of Buddhist monks burning themselves alive in the 1960s as a symbolic act, protesting against the repression of the population. This engagement in the social and political process became widely known through writings of the monk Thich Nhat Hanh, such as *The Lotus in the Sea of Fire*.

If China, with its strong Confucian traditions, did not really need the Buddhist religion to provide a leading civil ideology, the importation of Buddhism from India via Central Asia nevertheless presented a strong challenge to the incipient religious institutionalization of Daoism. Buddhism reached a high point of influence during the Tang dynasty, but it suffered greatly as a consequence of Daoist-inspired persecution in the mid-9th century. If the rivalry of the three religions had frequently led to strife and persecution, now of Daoism and now of Buddhism, there were also attempts to coordinate them. The classic model for this was the idea of the "three teachings" (*sān jiào*), which was most influentially set forth in a treatise ascribed to the first Ming emperor shortly after he had taken political control of the country. The compromise that he regarded as essential for political stability does not give a privileged place to Buddhism among the "three teachings," but it crystallized a model that even today implicitly underlies the law of a secular Chinese regime. Just as the Ming solution juxtaposed Confucianism on the one hand with Buddhism and Daoism on the other, so also now the state ideology of communism is the common reference point for the permitted religions: Buddhism, Daoism, Islam, Catholicism, and Protestant Christianity. Confucianism does not occur here as "religion," having been displaced by communist ideology as the normative backdrop, but Buddhism holds a clear position. The older "three teachings" concept implied the exclusion and suppression of the so-called heretical sects on the assumption that such minority sects lead to instability and are probably a cloak for agitation. This reasoning is therefore not a new communist invention, and observers need to consider contemporary cases carefully, for example, when new movements such as Falungong claim to be "Buddhist" and at the same time call for the overthrow of the government.

Buddhism was transmitted first to Korea and from there to Japan from the 6th century onward. If the dominant ideology in Korea up to the Japanese imperialist period was Confucianist, the postcolonial and postwar period has seen a resurgence of pride in the shamanistic traditions of that country, thus giving them a civil religious function. The political and social influence of Buddhism arises less from the traditional orders in Korea—strongly established though these are—and more from the newly founded lay movements such as the self-consciously modern Won Buddhism. Japanese history has seen immensely complicated variations of Buddhist influence via the imperial court on one hand and the power centers of major denominations on the other. While monks of Mount Hiei to the northeast of Kyōto occasionally led armed raids into this former capital, the millions of followers of the Shin Buddhist faith and others in the Nichirenite tradition were also a political force to be reckoned with and occasionally suffered dire consequences. In modern, postwar times, the influence of the Buddhism-inspired Kōmeitō ("Public Brightness Party") has been considerable but seems to be receding in the 21st century. Increasingly, lay Japanese Buddhist leaders have promoted a "peace" agenda, the Nichiren-oriented Risshō Kōsei-kai and the Sōka Gakkai being prominent in this respect. The strongly established Shin Buddhist denominations have led the way in reflecting on Japanese war responsibility and in educating their followers accordingly. In Sōtō Zen Buddhist circles, there has also been much heart searching while various influential Buddhist voices support the campaign to preserve Clause 9 of Japan's present constitution, prohibiting war as a means of settling international disputes.

Michael Pye
Philipps-Universität Marburg
Marburg, Germany

See also Confucianism; Monarchy; Religion

Further Readings

Dessí, U. (2008). *Ethics and society in contemporary Shin Buddhism.* Berlin, Germany: LIT Verlag.
Jones, K. (2003). *The new social face of Buddhism.* Boston: Wisdom.
Sarkisyanz, E. (1965). *Buddhist backgrounds of the Burmese revolution.* The Hague, Netherlands: Martinus Nijhoff.
Thich Nhat Hanh. (1967). *The lotus in the sea of fire.* London: SCM Press.
Zimmermann, M. (Ed.). (2006). *Buddhism and violence.* Lumbini, Nepal: Lumbini International Research Institute.

BUDGETING, RATIONAL MODELS

Budgeting refers to decisions about how much funding to provide for the programs, staffs, and infrastructures of government, as well as about the processes used to arrive at funding levels. In examining processes, budgeting theories explain how individuals who hold budget-writing positions choose funding levels for programs. A second class of theories explain how and why political elites try to manipulate budgets to foster their goals.

Rational theories of individuals responsible for choosing how much to fund budget lines offer bounded rationality explanations of these decisions.

Due to limits on human cognition, vast numbers of lines in the budget, and complex decision-making environments, budgeters' forecasts about how well their choices will facilitate their goals are imperfect. Despite these limitations, budgeters use decision-making strategies that help them foster their goals. Such strategies are boundedly rational because, although the choices and behavior resulting from such strategies are not optimal, budgeters' decisions still help them realize program goals. In this way, budgeters "satisfice"—they make decisions that are "good enough."

The first generation of boundedly rational budgeting theories was developed in the 1960s by Aaron Wildavsky, whose contributions to understanding budgeting continue to exert a profound influence on budgeting scholarship. This perspective, developed with respect to the federal government in the United States, emphasizes that individuals responsible for contributing to the budget must take an incremental approach to arrive at funding levels. Budgeters begin by examining the "base" for each budget line, which refers to the amount that was spent on the line in the previous year's budget. From this base, budgeters make small adjustments based on new information and limited by the availability of funds. For example, if program advocates can make an effective argument about why additional funds are needed, budgeters may increase funding levels slightly from the previous year. Nevertheless, even if budgets change from year to year, budgeters do not make drastic changes. Budgeters follow this approach out of necessity. Examining the budget holistically would involve the impossible task of investigating each budget line to ascertain the optimal volume of funds required to realize the goals embodied in the programs. Unfortunately, such an approach is not feasible given the vastness of the government and the resulting volume of decisions that must be made. By taking an incremental approach, however, budgeters arrive at funding levels for all budget lines, ensuring that programs continue serving their purposes even if neither the funding levels nor the achievement of program goals are optimal.

Additional research, however, observed that many budgeting decisions were not incremental, spurring refinement of Aaron Wildavksy's approach. John Padgett's serial judgment theory of budgeting stresses continuity with the incremental perspective in that budgeters' decisions are understood from a boundedly rational perspective. Budgeters cope with the volume and complexity of the decisions they face by starting with the level of funding from the previous year and satisficing to arrive at funding levels for the current budget. However, the heuristic employed by budgeters is not incremental choice. Rather than consciously choosing a funding amount that differs slightly from that disbursed previously, budgeters consider, one after another, alternatives for funding levels until they arrive at one that is acceptable. In this way, budgeters satisfice in that they do not try to discover optimal funding levels for programs. Critically, though, budgetary outputs are predicted to differ from outputs predicted by incremental approaches. Because budgeters move from alternative to alternative until they reach an outcome that is good enough, it is possible for them to arrive at funding levels that are substantially different from the amounts funded in the previous year's budget. Such dramatic changes can occur, for example, when budgeters' understandings of what is necessary from a programmatic standpoint leads them to pick a funding level that happens to be a significant departure from the previous year's base. Research focusing on the distribution of year-to-year changes in programs' funding levels in the context of U.S. government budgets revealed that this perspective explained better the distribution of year-to-year changes than an incremental approach did. By specifying more precisely the decision-making heuristics employed by budgeters, the serial judgment perspective improved on the incremental model while remaining anchored firmly to a realistically bounded rational view of human decision making.

More recent research by Frank Baumgartner and Bryan Jones integrates research on agenda setting with a boundedly rational approach to budgeting. Because the set of issues that they can attend to is small, budgeters ignore most information from the political environment. They do so even in the face of political mobilizations by coalitions favoring substantial funding changes. Under these circumstances, year-to-year changes in budgets are characterized by incremental shifts in funding levels. However, when groups in favor of change are successful in securing space for their concerns on the government's agenda, budgeters are forced to attend to the issues that these groups prioritize. In doing so, budgeters begin to consider alternative funding amounts for budget lines serially. Since past decisions related to the issues have resulted in incremental changes, political coalitions that have focused budgeters' attention on a specific budget line are likely to desire significant change in that line. Therefore, the funding levels that allow budgeters to satisfice in such instances are likely to constitute sizable increases/decreases. After all, substantial changes are the only ones likely to satisfy the political forces behind the mobilization. In this way, the heuristics that budgeters employ differ depending on whether advocates for change can secure space for their priorities on the agenda. If not, budgeters ignore information from such groups, leading to incremental changes. However, when agenda space is secured, serial judgment heuristics are triggered. What is more, the nature of the mobilization implies that only substantial changes in funding levels allow budgeters to allay the concerns of the coalitions—and, by extension, allow budgeters to satisfice.

Shifting gears, a separate tradition on government budgets asks whether political elites manipulate budgets to serve their goals. One line of such research asks whether ruling political parties stimulate the economy prior to elections in order to maintain their control of government and its decisions. This political budget cycle thesis was spurred by research conducted in the 1970s examining indicators of economic health in light of election cycles in the United States and in other major industrial democracies during the post–World War II era, observing that unemployment dipped and economic growth increased in years during which elections occurred. Naturally, the argument behind the thesis is that the ruling party prefers to remain in control of government. Therefore, it pursues policies that stimulate the economy to achieve economic growth prior to the election. Voters, the ruling party hopes, observe this performance and reward it by returning it to power.

Subsequent research on this hypothesis has produced mixed findings regarding the ability of ruling parties to pursue this strategy successfully. In recent research, James Alt and David Lassen suggest why: The ability of ruling parties to exploit their position is conditional on several factors. The first involves how transparent the government's fiscal policy decisions are to the public. Transparency allows the media and political opposition, and by extension the public, to observe the government's intentions, making it more difficult for the ruling party to try to create economic growth without being accused of doing so for its own benefit. The second factor is the level of polarization between parties. As policy conflict increases between the ruling party and its rival(s), the costs to the party of losing control of government to the opposition increases. These costs increase with policy conflict because the policies that the opposition (as the future ruling party) will adopt will be more offensive to the (present) ruling party for higher levels of policy disagreement between parties. In summary, while it may be feasible for the ruling party in government to rig the economy in order to maximize its chances of continuing in power, the circumstances under which this strategy is feasible may be limited.

Another line of research focusing on elite influence over budgets examines the participation of bureaucrats in shaping the budgets of their agencies. One theory advanced in the early 1970s held that bureaucrats were "budget maximizers" in that they wished to expand the size of their agency's budget as much as possible. What is more, bureaucrats are well positioned to do so because they possess more information about what is necessary to achieve the policy goals their agencies are responsible for meeting than elected officials who create budgets. This information asymmetry allows bureaucrats to receive substantial increases in their agencies' budgets.

Subsequent research in the tradition of the budget-shaping perspective by Patrick Dunleavy and others, however, has led to a substantial refinement of this view. In addition to noting that budgets are not characterized by the growth predicted by the maximization perspective, these critics note that, because an agency's budget is a collective good, bureaucrats face a collective action problem in advocating budget expansions. Moreover, because agencies' budgets are disaggregated, bureaucrats possess no guarantee that increases to the budget will accrue to their corner of the agency, undercutting the incentive to lobby. Bureaucrats also face varied incentives to lobby for expansions based on the type of agency they work for and their position within their agency. In particular, the highest ranking officials most capable of lobbying possess the least incentive to do so because, from a pecuniary standpoint, their salaries do not increase with budget increases. Instead of budget maximizing, the budget-shaping view argues that agency executives prefer to use the budget to construct an agency such that its staff perform professional functions characterized by high levels of discretion that foster creative solutions to problems. Executives in public agencies are thought to prioritize such "shaping" because individuals drawn to career advancement in such agencies are likely to value such a professional orientation. Additionally, and at any rate, laws forestall the ability of bureaucrats to capture pecuniary rewards from public service beyond their salaries.

Jason A. MacDonald
West Virginia University
Morgantown, West Virginia, United States

See also Agenda Setting; Political Economy; Public Budgeting; Rationality, Bounded

Further Readings

Alt, J. E., & Lassen, D. D. (2006). Transparency, political polarization, and political budget cycles in OECD countries. *American Journal of Political Science, 50,* 530–550.

Dunleavy, P. (1985). Bureaucrats, budgets, and the growth of the state: Reconstructing an instrumental model. *British Journal of Political Science, 15,* 299–328.

Jones, B. D., & Baumgartner, F. R. (2005). *The politics of attention: How government prioritizes problems.* Chicago: University of Chicago Press.

Padgett, J. F. (1980). Bounded rationality in budgetary research. *American Political Science Review, 74,* 354–372.

Wildavsky, A. (1964). *The politics of the budgetary process.* Boston: Little, Brown.

BUREAUCRACY

The term *bureaucracy* denotes a particular form of organization that is complex and differentiated, has defined rules and procedures, and is subject to a command-and-control system of hierarchical authority. This entry discusses the nature and functions of bureaucracy and the bureaucratic state, as well as the problems and deficiencies associated with it.

Definition and Characteristics

The illustrious German sociologist Max Weber noted several characteristics of the bureaucratic form of organization that defined it as an ideal type.

First, bureaucracy derives its authority through law. It depends neither on traditional nor on personal modes of authority. It is instead based on a form of legal-rational authority that defines the foundations of the organization's jurisdiction and its procedures of operation.

Second, bureaucracy has differentiated offices or units that have specialized competencies and jurisdictions. This aspect of bureaucracy provides it with some obvious advantages, including the ability to bring technical expertise and experience to bear on particular aspects of problems. It also raises some difficulties, however, in coordination of tasks, and it sometimes creates conflicting perspectives as to what the task is or how a problem should be defined.

Third, bureaucracy is characterized by an internally hierarchical system of authority required to bring its distinctive parts together as part of a unified system of coordination. The extent to which this hierarchical control system can actually be put in place is almost always imperfect, though the extent to which it can be consummated rests on many factors other than coercion. These include leadership that obtains support from all of the organization's distinctive offices and units, clarity of purpose from organizational leadership, and careful assessments of feasible paths to ends.

Fourth, bureaucracy is also characterized by rules and procedures that govern its internal functions. It has a memory and a set of procedures for dealing with recurring matters. Weber referred to this as the organization's "files," and this storage of experience and written rules anchors bureaucracy in the legal-rational system of authority. The files provide precedents and also produce standard operating procedures (SOPs) that enable bureaucracies to become more proficient in responding to problems that can be readily identified and are within the organizational repertoire. The more ambiguous or multifaceted the problem, however, the more complicated will be the sorting process within a bureaucratic organization and the more likely, therefore, that problems are recognized and defined in ways reflecting the complex differentiation of bureaucratic organizations. This means that there are likely to be many different perspectives rather than a unified one.

Fifth, bureaucracy is typically characterized by a full-time corps of officials. In governments, these officials are known as civil servants; they represent the continuity of the organizations of government, their missions and functions, and the antecedent ongoing commitments of government. Similarly, in private enterprise and even in nonprofit organizations (nongovernmental organizations), full-time managers and subordinates provide the basic maintenance functions within an organization, negotiate its external environment, and organize its production and marketing functions.

Realities are often more complex than this ideal-type description of bureaucracy. Sometimes bureaucratic forms exist but without the reputed benefits of expertise and efficient allocations of attention. Bureaucratic pathologies may run deep reflecting both environmental and cultural influences, shortages of resources, and incentives for suboptimal performance. These pathologies are especially notable in less developed countries. Consequently, to understand the topic better, this entry addresses the following questions: Why do we have bureaucracy? What are the relationships with capitalism and market? What is the bureaucratic state and what are its relationships with democracy? What are the internal and external pathologies of bureaucratic behavior?

Why Do We Have Bureaucracy?

In Weber's ideal-type description of bureaucracy, he noted that bureaucracy was the technically most proficient and efficient form of organization. Far from the negative stereotypes that have arisen in the contemporary era from pathologies associated with bureaucracy, Weber observed that bureaucratic organization was a necessary concomitant condition for the rise of capitalism based on economic efficiency. The rise of the modern public stock corporation separated management and ownership making the managers extremely important to the direction and guidance of the firm.

Capitalism necessitated legal-rational forms of authority. The modern corporation is itself a legal entity. Unlike a person or even a partnership, corporations have limited liability for their officers. Corporations and other types of firms require legally grounded rules in a capitalist system. Disputes must ultimately be settled through juridical means, and agreements must be carefully and legally delineated so that there is a record—what Weber referred to as "the files"—or a recorded precedent. Impersonal authority is central to a system of rational-legal authority, and bureaucracy is the embodiment of an organization defined by rules of procedure and responsive to external legal authority.

Above all, complexity requires organizational differentiation, specialization, and limited jurisdiction coordinated through a hierarchical system of authority. These crucial aspects of bureaucracy not only allow for greater efficiency, economy of effort, and expertise but also lead to problems of coordination.

Bureaucracy and Markets

On the face of it, bureaucracy and capitalism seem to live in an uneasy coexistence. One is structured by authority and legal rules, while the other in its pure form reflects an uninhibited free exchange of goods and services mediated by the price mechanism. The reality, however, is much more complicated. Sellers are not necessarily numerous, and where possible, they seek to control markets and thus influence, if not dictate, prices. Public stock corporations' interests lie in increasing the value of their stock. Rarely are consumers fully armed with perfect information as classical theories of economics assume. The rise of governmental regulation of markets and marketing practices and of the safety of products provides a means of short-circuiting the information needs that consumers by themselves would be unlikely to obtain. As a consequence, the rise, and, to some degree, the inevitable excesses of the profit motive lead to the necessity of regulation and the growth of governmental bureaucracy. Newly capitalistic systems, in fact, often suffer from a rapid expansion of production without adequate means of regulating the quality or safety of what is being produced. Such mechanisms tend to evolve in developed capitalist systems in order to counter the adverse impacts of unregulated capitalism.

The ascendency of capitalism and the modern corporation roughly parallels the rise of professionalism in government and the development of the modern state, characterized by governmental bureaucracy. Ironically, while it is sometimes claimed that capitalism rests on the free exchange of goods and services through the price mechanism and is thus antithetical to bureaucracy, the reality is that modern governments, through their bureaucracies, pave the way for markets to function by providing infrastructure, regulation of markets, transparency in market processes, and regulation of the relationships between private and public interests. In reality, states and markets have been more often collaborative than in conflict. Indeed, it is hard to imagine the effective functioning of markets without the effective functioning of government. The complexity of the modern corporation and the complexity of modern government have grown together more symbiotically than antagonistically.

The Bureaucratic State

It is hard to imagine modern government without bureaucratic capability. Weak states lack capability. And weak states are often characterized as ones affected by civil strife, lacking effective records, influenced by corruption, and unable to bring services to their populations, especially those that are remote from their capital cities. They are also typically incapable of collecting revenues in an impartial fashion. Cronyism runs amok. A professional corps of civil servants is usually absent and government is most often turned into a spoils system of tribal winners—or, as a means of keeping peace among contending factions, sometimes awarding some of the spoils to each of them. Weak states are inept precisely because they lack the desiderata of professional bureaucratic organization and a professional civil service and the resources to ensure their ability to function. Consequently, modern government and bureaucracy are nearly interchangeable. Continuity, specialized expertise, a professional corps of civil servants, the rule of law and impersonal authority, and documented records are the essence of bureaucracy and of modern states capable of governing, providing services, and collecting revenues.

For the most part, the bureaucratic state largely preceded the arrival of democratic government. It served the interests of a sovereign crown and typically led to the crown's extension of its authority and, eventually, to what we know of today as the modern state. As a result, the state bureaucracy often has been seen as a conservatizing force in society—the consolidator of existing authority rather than a challenger to it. With the growth of mass enfranchisement, the bureaucracy was often viewed as the upholder of the existing regime and a potential threat to democracy. Its power rests on its continuity and specialized expertise and its permanence and durability, while the temporarily mandated leaders of government are often in a position of dependence on the bureaucracy. In a frequently invoked phrase, are the experts on tap or on top?

One major exception to the bureaucracy–democracy sequence was the United States where mass enfranchisement among White males preceded the development of a professionalized bureaucracy. Even after a professional civil service emerged, the federal government had few responsibilities until the 1930s, when the Franklin Roosevelt administration extended the federal role into the realms of economic and business regulation, the provision of social insurance, and, ultimately, the welfare state.

There has been much discussion across many developed societies about transforming bureaucracy, lightening its load, and moving toward market-oriented incentives, but especially in the United States, the bureaucracy has been deeply controversial. There, the identification of bureaucracy with "big government" and with controversial social reforms makes it both a point of partisan division and an object of derision.

The essential point is that while a facade of bureaucracy is no guarantee of effective government, it is virtually impossible to imagine effective government without bureaucracy. That, however, does not mean that public bureaucracy will be without controversy or demands for reform to which it has been subjected especially since the 1980s. Nor does it mean that the bureaucratic organizational form is without contradictions or that it can lead to certain pathologies.

Bureaucratic Pathologies

Despite the advantages conferred by bureaucratic organization, pathologies result from the internal contradictions of this form of organization. These we will call internal pathologies. Another set of pathologies results from the ways in which bureaucracies relate to their environment. These we will call external pathologies.

Internal Pathologies

Analysts, as well as antagonists, of bureaucracy have pointed to a variety of problems stemming from the logic of bureaucratic organization. These are presented in no particular order of importance.

One pathology stems from the very system of impersonal authority and legal rationality on which bureaucratic organization is based. As a rule-driven form of organization, bureaucracies must defend decisions based on legal rulings and precedents. Consequently, bureaucracies may be short on compassion, empathy, or extenuating circumstances of case equity. Uniform procedures rather than case equity are a hallmark of bureaucratic organization. A frequent criticism of this formalism in bureaucracy is characterized by accusations directed to "faceless," "nameless," or "unfeeling" bureaucrats whose decisions defy "common sense." Actually, some studies of bureaucrats dealing with clients indicate that bureaucrats themselves often try to provide sympathy, guidance, and solace within the constraints of the law. But law is what guides bureaucracy, and this particular pathology, in reality, is a critique of the rule of law. By contrast, personal relationships often govern politics, and these relationships, in turn, generate charges of favoritism.

A second pathology stems from the command and control system of hierarchy in bureaucracies. Hierarchical authority is necessary to coordinate the various specialized jurisdictions as well as to provide an ultimate point of organizational accountability. Still, there is a continuous tension in all complex organizations between strict control from the top and devolution of authority to those at more operational levels within an organization. Paradoxically, control from the top downward produces more definitive accountability while also stymieing an organization's ability to adapt to problems it encounters on the ground. It also may encourage risk-averse behavior and a desire on the part of subordinates to keep chastening news away from their superiors. Studies of military units observe that noncommissioned officers (the bottom rung of authority) are crucial to adaptability and to organizational effectiveness. How to apportion and allocate authority in bureaucratic organizations in ways that keep it from flying apart and thus adhering to no common identity remains unresolved. At the same time, how to imprint leadership direction without imposing stultifying restrictions on operational units is also unclear. Organizations need to be responsive and adaptive to their environments but frequently cannot be if control from the top is excessively restrictive. Obviously, balance is needed, but it is unclear, certainly in the abstract, as to exactly where that balance should be struck. Without sufficient control, unaccountable behaviors may become rampant—a potential issue especially for organizations with many "street-level" bureaucrats. But too much control is likely to lead to cautious, perhaps even timid, behaviors on the part of those responsible for hands-on operations.

A third pathology of bureaucratic behavior stems from the differentiated and specialized jurisdictional character of bureaucratic organization. Especially when individuals have spent much of their careers in a given unit, they are likely to look at and define problems within the definitional context of their units. A number of studies show that each unit tends to see the "face of a problem" in a way consistent with their unit's responsibilities and the subculture of their unit. Given the tendency to search for problem definition and responses in a cognitively economical fashion, responses are apt to rely deeply on how specific organizational units define and codify responses and perpetuate them. Does this produce a cacophony of voices or a necessary diversity of perspectives? Again, a problem of finding the ever-elusive right balance appears. Organizational subunit responses are not random. In fact, they are typically predictable. But the various subunits' responses cannot all be equally satisfied, which is, of course, where leadership comes in. Differentiation and the codification and subcultural practices that stem from the distinctive component parts of bureaucratic organization provide predictability but not synthesis, which then becomes the responsibility of central organizational leadership.

Studies, even of crises, show how distinctive "the face of the problem" will appear to different organizational units in accordance with their distinct missions and their core cultures, technologies, and perspectives. Can unity and clarity of direction be woven against the backdrop of these distinctive perspectives? And can central organizational leadership coax the losers in bureaucratic battles to cooperate rather than to continue the struggle covertly?

A fourth potential pathology of bureaucratic organization is the reliance on precedent. One of the most important characteristics of bureaucracy is that it provides for continuity. Critical to continuity is the recording of precedents and the use of precedents as a platform through which to assess current issues. Precedent serves to economize on cognitive energy. Basing current choices on precedents reduces the costs of search for organizations. However, reliance on precedents may lead to a highly conservative outlook temperamentally, one that may be more wedded to the past than to the future. Bureaucracy tends to be thought of as a lumbering status quo form of organization. That is due in no small part to the emphasis on following rules and procedures and the reliance on precedents set by the organization. Ultimately, bureaucracies are expert at routinizing the previously unexpected or contingent. But once this is done, continuity with precedents becomes part of the operating modality and departures that may be required for adaptability, equity, or even sensibility become harder to achieve. Bureaucracy provides ballast and stability but by itself is unlikely to provide fresh direction.

A fifth possible pathology lies in the intersection between the individual interests of bureaucrats and collective well-being. Some analysis based on economic models, for example, suggests that bureaucrats have a stake in the growth of their agencies' funding as they themselves are apt to benefit from the growth of revenue. A more compelling interpretation, though, would propose that where the fundamental technology of an organizational unit is invested in significantly by the bureaucrats of that unit, changes in the core technology puts at risk those whose training has been part of a prior technology. The more intensive and specialized that training, the more resistance there may be to alterations in it that would threaten the status of the current technology. The usual cliché about this is that all military forces are prepared to fight the previous war. This bit of conventional wisdom may well be overstated but there is a fundamental truth behind it, and that is that no one likes to be displaced by a newer or different technology that renders one's own training obsolete. The human being inside the machinery of complex bureaucratic organization, therefore, needs to be taken into account.

In summary, the virtues of bureaucratic organization also contain the seeds of its deficiencies. Continuity, predictability, legality, precedent, and specialized expertise, accordingly, may lead to difficulties in adapting to change or recognizing novelty, difficulties in applying case equity, and problems of coordination. Hierarchical authority also can lead to a lack of sensitivity to problems on the ground. Yet a lack of hierarchical control may lead to a lack of accountability. Because these problems are built into a system of law and impersonal authority, there are no set formulas for resolving them. But they remain problems inherent in bureaucracy.

External Pathologies

Despite our dependence on bureaucracy to provide orderly processes, the virtues of bureaucracy tend to be hidden to the unschooled eye, whereas its deficiencies seem to be obvious. One of these defects lies in the rule boundedness of bureaucracy resulting in the well-known "red tape" problem. In contrast to the realm of personal authority where one may petition to have a problem fixed through political means, procedural fairness and the accountability of bureaucracy to legal and auditing requirements inevitably produces large amounts of paperwork. The objective is to reduce the improper use of funds and procedural irregularities and to ensure a record of transactions. Few people, even bureaucrats, are enthralled by red tape. Most are frustrated by it. But red tape is the product of demands for accountability and procedural regularity, especially in the public sphere. A bureaucracy notably responsive to its clienteles might be regarded as showing favoritism and evading lawful procedures. The reality is that while we extol the rule of law in the abstract, we often prefer personal relationships concretely. That is, we often want a "fixer" to deal with the complications that regularity of procedures demands. Everyone wishes to cut through the maze of red tape to get what they want, but they also want to have procedures in place to deter others from getting what is improper.

Public bureaucracies particularly are expected to be both accountable and responsive. These terms, however, are remarkably elusive. *Accountable* to what and *responsive* to whom are questions that immediately arise. There appears to be no definitive agreement as to what precisely these words mean. Often, they are used interchangeably. If, however, accountability is taken to mean being answerable to a legally grounded authorization—a constitution, a sovereign, a law, or a fiduciary responsibility—and if responsiveness is taken to mean acceptance of the demands of temporary political authorities or of agency clienteles, one can see how these ideas may come into conflict. The great dilemma is for bureaucracies to be accountable and responsive up to a point—a point, however, that rarely can or even ought to be satisfied fully. An overly responsive bureaucracy will no longer be seen as neutral, fair, or judicious. A bureaucracy, however, that is unresponsive either has succeeded in displacing the elected or, at the other extreme, has lost its political relevance. Bureaucrats need some independence from the current political authorities if they are to function properly, but the line demarcating independence and unresponsiveness is very murky indeed. To find the appropriate balance is as difficult as it is necessary. Studies of senior civil servants in many developed countries indicate that they are sensitive to their political context but are frequently concerned about the proper equilibrium between neutral competence and accountability on the one hand and political/policy responsiveness to the current authorities on the other.

From a political perspective, another criticism of bureaucracy, especially as a regulator of business is that it is a "deadweight" on the functioning of markets by imposing costs on firms that add to prices. While regulation by bureaucrats is the product of political decisions made by duly authorized governments, free market proponents believe that bureaucratic regulation diminishes economic dynamism and efficiency. Each new regulation drives up the cost of doing business and restricts the options that businesses may have available to them. When the regulatory hand is made lighter, however, risks often arise that result in crises or avoidable tragic outcomes, and the call for regulation is renewed. The proper balance between regulation and business flexibility, as with other relationships between desirable but contradictory values, is hard to define, and the proper instrumentation for overseeing and enforcing regulatory behavior is equally difficult to pinpoint.

Finally, a significant criticism of bureaucracy is that its incentive structures are all wrong. Civil servants are insulated through tenure. This insulation may lead to behavior that provides little incentive to promote efficiencies that could disturb the status quo since there is little reward for doing so. A contrast is frequently drawn with private enterprise where continuous innovation is believed to be necessary for a firm's health and growth prospects. According to this perspective, the business model is presumed to be the form of organization worth emulating, though, of course, most large-scale businesses are also bureaucracies. Administrative-reform efforts beginning in the 1980s, falling under the rubric of "new public management," essentially emphasized the business model, emphasizing the off-loading of many government functions to the private sector and providing managerial incentives to improve performance in a scaled-down public sector. The assumption was that what worked in the market place would work equally well in government. Freeing up the managers and holding them accountable for performance was one reform pathway. Creating agencies and managers with limited-term contracts was another. In at least one case, tenure itself was eliminated in the public sector.

The results of this emphasis on making government more like business are still uncertain, and the movement toward providing incentives for promoting efficiency and improved performance may rest on an as yet unconfirmed assumption that officials in the public sphere have similar motivations as their counterparts in the private sphere. In fact, what evidence there is on this matter suggests otherwise. In the business model critique of bureaucracy, performance should be based less on adherence to rules and more on results. A lessening of rules-based administration, one can imagine, would be open to extensive challenges, and a fundamental search for consistent application of law would hardly disappear. Emphasizing results, of course, requires agreement on what the results should be and on the metrics required to assess performance. It is likely that some agency missions may not be easily susceptible to clear metrics. It is also likely that some agencies will adjust the definition of their missions in ways that allow for favorable metrics even if they are peripheral to their broader, if ill-defined, missions.

Living With Bureaucracy

Not all organizations are bureaucratic and, indeed, some of the most innovative ones in the private sector tend to be flat and intermingled to some degree in their functions. However, the growth of bureaucracy reflects modernity and a system ruled by legal authority and by differentiated function. That produces annoyances such as the heavy emphasis on procedures, rules, and specialized offices whose understandings may not be fully integrated with one another. Nevertheless, bureaucracy is able to provide continuity in government, consistency in application, proficiency through specialized and limited jurisdictions, and professionalism through a full-time civil service based on merit.

The limits and critiques of bureaucracy are notable. Its virtues are also often its vices. Balancing the contradictory tensions of the bureaucratic form of organization presents a continuing challenge. There is also no doubt that much of the criticism of bureaucracy is distinctly political and has little to do with bureaucracy itself and more to do with views about the appropriate scope of government. Finally, to appreciate the advantages of bureaucracy, it is useful to contemplate life without it. It would be a world without continuity, without consistency, and with a shortage of competence. We may not like bureaucracy, but we would assuredly dislike a world without it even more.

Bert A. Rockman
Purdue University
West Lafayette, Indiana, United States

See also Administration; Administration Theory; Bureaucracy, Rational Choice Models; Bureaucracy, Street-Level; Capitalism; Democracy, Theories of; Effectiveness, Bureaucratic; Market Economy

Further Readings

Aberbach, J. D., & Rockman, B. A. (2000). *In the web of politics: Three decades of the US federal executive*. Washington, DC: Brookings Institution Press.

Clapp, P. A., & Halperin, M. H. (with Kanter, A.). (2007). *Bureaucratic politics and foreign policy* (2nd ed.). Washington, DC: Brookings Institution Press.

Gerth, H. H., & Mills, C. W. (Eds. & Trans.). (2001). *From Max Weber: Essays in sociology*. London: Routledge. (Original work published 1946)

Kaufman, H. (1977). *Red tape: Its origins, uses, and abuses*. Washington, DC: Brookings Institution Press.

March, J. G., & Simon, H. A. (1993). *Organizations* (2nd ed.). Cambridge, UK: Blackwell. (Original work published 1958)

Peters, B. G. (2001). *The politics of bureaucracy* (5th ed.). London: Routledge.

Simon, H. A. (1997). *Administrative behavior: A study of decision-making processes in administrative organizations* (4th ed.). New York: Free Press. (Original work published 1947)

Thompson, V. (1975). *Without sympathy or enthusiasm: The problem of administrative compassion*. Tuscaloosa: University of Alabama Press.

Wilson, J. Q. (1989). *Bureaucracy: What government agencies do and why they do it*. New York: Basic Books.

Bureaucracy, Rational Choice Models

This entry discusses rational choice models of the bureaucracy. It first defines what it means for a theoretical model of the bureaucracy to employ rational choice and discusses two major applications of these models. It then discusses the basic assumptions underlying most models of the bureaucracy. Finally, it presents major themes and findings in the literature and discusses how these are related to modeling assumptions.

Rational choice models of the bureaucracy operate under the assumption that all relevant actors are rational. All players in the model—the government, bureaucrats, and private sector entities affected by bureaucratic decisions—have well-defined preference orderings over possible outcomes and act in a manner that ensures that their preferred outcomes are realized.

To facilitate analysis, scholars in this tradition typically assume that players' preferences satisfy several properties. It is generally assumed that players' preferences can be represented by a *utility function*, such that a player prefers outcome A to outcome B if and only if that player's utility from outcome A is greater than that from outcome B. It is often further assumed that players have consistent preferences regarding risk, so that preferences over uncertain outcomes may be represented by *expected utilities*. Many models further assume that players react to new information according to Bayes's rule (i.e., how one should change one's existing beliefs in light of new evidence).

Rational choice models of the bureaucracy tend to be game theoretic. All players are assumed to be strategic in their interactions—each player considers the strategies of all other players in attempting to reach his or her preferred outcome. Analysts derive equilibrium solutions to these models, where an equilibrium is defined as a set of strategies wherein each agent's strategy is a best response to all other agents' strategies, and agents' beliefs are consistent with those strategies.

The rational choice literature on bureaucracy has primarily focused on two, interrelated, problems: delegation and special-interest influence. Models of delegation involve a government granting authority to a bureaucratic agent. The crucial problem for the government lies in the fact that the bureaucracy may not perfectly share its preferences. The government must, therefore, oversee bureaucratic decisions and provide incentives to ensure that the bureaucracy acts in the desired manner. This problem closely relates to the branch of microeconomics known as principal–agent (or contract) theory, which examines a similar problem between firms and their employees.

Models concerned with the influence of special interests focus on the interaction between the bureaucracy and the private sector. Private sector agents may wish to manipulate the behavior of bureaucrats. The mechanisms to do so may involve bribery, appeals to bodies charged with bureaucratic oversight, or the provision of jobs to former members of the bureaucracy (following the principle of the "revolving door). Such models may also relate to delegation, as the government may wish to curtail special-interest influence and will conduct oversight accordingly.

Model Assumptions

The Preferences of Bureaucrats

Since rational choice models of bureaucracies assume that bureaucrats act to achieve their most preferred outcomes, these models critically depend on assumptions about what bureaucrats prefer. Preferences may be assumed to be *primitive*, or they may be *induced*. Primitive preferences describe outcomes bureaucrats inherently enjoy. Induced preferences are defined by actions. Bureaucrats are driven to prefer certain actions as, in equilibrium, these lead to preferable outcomes.

The rational choice literature on bureaucracies has generally defined bureaucrats' preferences in one of three ways. William Niskanen introduced the assumption that bureaucrats attempt to maximize their agencies' budget allocations. This preference is treated as primitive for modeling purposes. But Niskanen argues that the preference for budget maximization results from an implicit theory of the interaction between bureaucrats and the government. An individual bureaucrat is presumed to attempt to maximize his or her salary (both currently and in the future) and prestige. Both salary and prestige depend, in part, on the budget of the bureaucrat's agency. Since the bureaucrat's salary and prestige are assumed to be improving in the budget allocated to the bureaucrat's agency, the bureaucrat acts to increase this allocation.

Other models of bureaucracy work with assumptions regarding bureaucrats' primitive economic preferences. Like other economic agents, bureaucrats attempt to maximize consumption and leisure. Such models may be used to *derive* induced preferences over actions. For instance, *career concerns* models predict that bureaucrats act to maximize their reputation in time t because this reputation affects their career prospects, and thus consumption, in time $t + 1$. This result is an equilibrium prediction from the model—an author of a career concerns model need not assume that bureaucrats have primitive preferences regarding their reputation.

Finally, many models of bureaucratic politics treat bureaucrats (and governments) as having preferences over policy outcomes—that is, ideological preferences. Ideological preferences are generally treated as primitive. Individual bureaucrats may have inherent preferences over policy outcomes—a reflection of their individual political ideology.

Alternatively, ideological preferences may be treated as induced by an unmodeled interaction between bureaucrats and other entities. For instance, many models of the American bureaucracy assume that agency preferences reflect those of the president. Because the president is able to appoint and remove many agency heads, it is assumed that bureaucrats will behave *as if* they share the president's ideology.

Information

The information available or not available to players shapes their strategies to achieve their preferred outcomes. In models of delegation, it is typical to assume that bureaucratic agencies have an informational advantage vis-à-vis the government. The government delegates rule-making authority to agencies because bureaucrats have technical knowledge and experience that are not available to the average member of government. Governments may be less able than bureaucracies to design policies to achieve desired outcomes.

Other forms of informational asymmetries may also be incorporated into models of the bureaucracy. For instance, if the government cannot perfectly observe the skill or ideology of potential bureaucrats, it is faced with a problem of *adverse selection*. In models of adverse selection, the government must conduct oversight and provide wages keeping in mind the type of agents these incentives will attract. For instance, if some possible recruits are willing to engage in corruption, while others are not, sufficiently low wages may attract corruptible recruits who are better able to supplement their salaries with illicit income.

If the actions of bureaucrats in office are only imperfectly observable, then governments are faced with a problem of *moral hazard*. It is important to note that moral hazard is only an issue if outcomes are also imperfectly observable or are imperfectly determined by bureaucrats' actions. If bureaucrats' actions perfectly map into observable outcomes, then the government can accurately deduce bureaucrats' behavior based on these outcomes. However, if this is not the case, then the government must conduct oversight in a manner that minimizes the costs of moral hazard. These costs arise because the government is not able to optimally reward and sanction agent behavior. Instead, it can only respond to outcomes that imperfectly reflect this behavior.

Finally, all players may lack information over the future composition of the government due to political uncertainty. Members of the current government may be replaced by elections, coalition realignments, retirement, or death. As new individuals enter government, the ideological preferences of the government may change. All players in the model are therefore uncertain of future government policy preferences.

Government and Bureaucratic Structure

The final elements in a model of bureaucratic politics are assumptions about the number and composition of players. Models of delegation must include, at least, a unitary government and a single representative bureaucratic agent. Similarly, models of interest-group influence require at least a representative bureaucrat and a single special-interest group. But more sophisticated models may incorporate more realistic assumptions about the number of players and their interactions—with important effects on equilibrium outcomes.

For instance, models may vary the number of players considered to be in the government. Some models may incorporate the existence of multiple legislators or the separation of powers between branches of government. Such a model is said to include *multiple principals*.

Similarly, bureaucracies may be modeled as consisting of multiple bureaucrats or multiple agencies. Models incorporating more than one bureaucrat are said to feature *multiple agents*. Bureaucrats or agencies may be redundant—that is, assigned to similar or identical tasks. Alternatively, some agencies may be granted oversight authority over others.

Finally, all models of special-interest influence—and some models of delegation—incorporate private sector agents. These may be one or more firms seeking to maximize profits, one or more citizens or groups thereof with ideological preferences, or groups of consumers. The number and composition of these interests affect their interactions with the government and the bureaucracy.

Models of Delegation

As noted above, models of delegation involve a government's decision to grant authority to a bureaucratic agent, who may not perfectly share the government's preferences. These models are used to derive predictions about which bureaucrats will be appointed, how they will perform in office, what forms of oversight will be employed, and how much discretion they will be granted. These predictions will depend, in part, on assumptions regarding the preferences of—and information available to—all players, as well as the number and composition of players in the model.

Appointment

Models of appointment attempt to deduce which candidates will be selected for bureaucratic office, where candidates vary on some dimension. Typically, models of appointment assume that candidates for bureaucratic office have primitive preferences over policy, and candidates vary in their preferred policies.

In most settings, models of bureaucratic appointment predict that governments will appoint officials whose primitive ideological preferences are most proximate to the government's own. This finding is known as the Ally principle. The logic of this finding is straightforward: If a bureaucrat is granted discretion to implement policy, his or her policy choice will be dictated by his or her preferences. The closer the bureaucrat's ideological preference to that of the government, the more the government can expect to benefit from the bureaucrat's policy choice. This principle relates closely to the literature on attempts to politicize the bureaucracy.

However, the Ally principle may be violated under conditions of uncertainty. Assume now that candidates for office vary not only in their ideological preferences but also in their technical knowledge. More precisely, assume that some candidates are better informed of the mapping from policies into outcomes than are others. In this instance, the government may prefer to appoint a more knowledgeable candidate with preferences more removed from its own over an ill-informed candidate with proximate preferences. The ideologically proximate candidate may be unable to ensure that the policies he or she would adopt in office would, in expectation, lead to outcomes that the government prefers over policies that the more ideologically distant—but more competent—candidate would adopt.

The Ally principle may also fail to hold in settings where bureaucratic appointees interact with other players in formulating policies. For instance, a government may prefer to appoint a central banker who is more inflation averse than the government itself. This result holds because policy outcomes—here inflation and unemployment—reflect the interaction between the central banker and the public. Central bankers cannot commit to implement policies they do not prefer. Therefore, if the central banker is not inflation averse, the public will expect him or her to attempt to generate unexpected inflation—thus reducing unemployment. These inflationary expectations produce inflation—and no reduction in unemployment—in equilibrium. However, if the central banker is inflation averse, the public will expect inflation to remain low—a more favorable outcome for the government.

Bureaucratic Slack

A second major concern of models of delegation is the danger of bureaucratic slack. Bureaucrats are said to be shirking—yielding slack—when they exert little or no effort in their duties. Slack is typically a feature of models wherein bureaucrats are assumed to have economic preferences—they maximize consumption and leisure. Slack thus occurs when bureaucrats face insufficient incentives to forgo leisure.

Questions of bureaucratic effort are intimately linked to the power of the incentives bureaucrats are offered. Incentives are said to be *high powered* when rewards and punishments are heavily conditioned on outcomes. They are *low powered* when rewards and punishments are largely independent of outcomes. Low-powered incentives tend to increase slack.

The degree of bureaucratic slack will also depend, in part, on the structure of information in the model. In particular, moral hazard will tend to increase shirking. The rationale behind this result is intuitive: If the government is unable to determine the degree of effort exerted by bureaucrats, it will be difficult to reward or punish this effort. It is therefore difficult to encourage bureaucrats not to shirk.

Slack will also be affected by the number of players with authority over the bureaucracy. Models with multiple principals will tend to have lower-powered incentives and greater slack than models with a single principal. Slack will also tend to increase as the interests of the principals diverge. So, for instance, one might expect greater bureaucratic slack to occur when powers are separated between branches of government and when different parties dominate each branch.

The presence of multiple government principals may produce bureaucratic slack for one of two reasons. First, multiple principals may all try to induce bureaucrats to behave in their preferred manner. If principals have different interests, these incentives will tend to conflict, offsetting one another. This type of model is most likely to make sense when depicting a lower level bureaucrat with multiple overseers.

Alternatively, multiple agents within the government (e.g., legislators) may need to coordinate to punish or reward a bureaucrat for his or her performance. As the number of legislators with conflicting interests rises, it grows less likely that a reward or punishment will be agreed on.

Slack will also be affected by the presence of multiple bureaucratic agents in a model. Redundancy—the assignment of an identical task to more than one bureaucrat—will tend to increase slack when bureaucrats only care whether one member of a team performs a task successfully. Such a situation may result if, for instance, multiple police officers are assigned to a given case and all share credit if the case is successfully solved. In such a situation, each officer prefers to shirk while the others work on the case. He or she then might enjoy leisure time and benefit from the rest of the team's efforts. Since all officers have the same preference for slacking, no officer will exert the optimal amount of effort in solving the case. As a result of this problem, the probability that at least one member of the team solves the case may not rise with the number of officers assigned.

However, in situations characterized by moral hazard, redundancy may reduce slack. In such a situation, the government may compare the performance of one bureaucrat or agency with that of other bureaucrats or agencies charged with the same task. This comparison will allow the government to deduce better the level of effort exerted by the agency and reward or punish it accordingly.

Low-level bureaucrats may also slack in models where the bureaucracy assumes a tiered structure, such that some agents are given oversight authority over others. In such a system, low-level bureaucrats will tend to face low-powered incentives. Were they to face high-powered incentives, low-level bureaucrats may attempt to bribe their overseers to ignore poor performance. As the power of their incentives increases, low-level bureaucrats have greater reason to collude with their overseers and are thus willing to pay larger bribes. The government will thus optimally reduce the power of incentives faced by low-level bureaucrats so as to deter such collusion.

Bureaucratic Drift

Bureaucratic drift refers to the phenomenon wherein bureaucrats adopt policies that differ from those preferred by the executive or legislature. Drift is therefore typically a concern in models in which bureaucrats have preferences over policy. Drift results when bureaucrats are given discretion to implement their preferred policy outcomes.

Drift is typically controlled through one of two mechanisms: *ex ante* limits on bureaucratic discretion or *ex post* review of bureaucratic decisions. The former refers to legislative limits that prohibit certain policy positions from ever being adopted. The latter refers to the ability of other government entities (the legislature, the courts, etc.) to review and overturn bureaucratic decisions. These two mechanisms act as substitutes in preventing bureaucratic drift—as *ex post* review becomes more difficult, *ex ante* discretion is likely to decline.

The extent of bureaucratic drift will be affected by the extent of the informational asymmetry between the bureaucracy and the government. The more technically knowledgeable the bureaucracy, relative to the government, the fewer *ex ante* limits on discretion. As levels of discretion expand, the bureaucracy has greater latitude to implement its own ideological preferences.

Adverse selection may also affect the degree of policy drift. A provocative recent paper demonstrates that the degree of policy discretion afforded to bureaucrats may affect their willingness to remain in office and to acquire policy relevant expertise. Policy-motivated bureaucrats will be more likely to invest in acquiring expertise if they are granted discretion. To the extent that discretion is limited, policy-motivated bureaucrats will value office and expertise less. Such officials will be less willing to exert effort to arrive at correct decisions when discretion is limited. In other words, policy drift may be a necessary cost for an informed bureaucracy.

The number of government principals may also play a role in bureaucratic drift. If these principals must coordinate to overturn bureaucratic decisions—as in a legislature—their ability to exercise *ex post* review may be limited. The greater the number of principals and the more conflicted their interests, the less likely *ex post* review will be exercised. To combat the resultant danger of drift, *ex ante* restrictions on bureaucratic discretion will tend to rise.

Finally, the threat of changes to the governing coalition may affect drift. Suppose that the government at time t fears that it will be replaced by an alternative government with divergent preferences in time $t + 1$. To the extent possible, such a government prefers to lock in current policy choices, limiting the policy discretion of successor governments. One means of so limiting policy discretion is to adopt administrative procedures that limit the oversight authority of future governments. Since future governments may use their oversight powers to force through undesirable policy changes, the government at time t will seek to insulate the bureaucracy from such oversight.

Models of Special Interests

Models of the interaction between special interests and the bureaucracy are centrally concerned with when, and whether, special interests will be able to influence bureaucratic decisions. Influence may emerge through a number of mechanisms: bribery, offers of future employment, and appeals to the government to review bureaucratic decisions. Influence might also be exercised through more subtle mechanisms—for instance, through threatening appeals to overseers. Special-interest influence may affect the agency relationship between the government and the bureaucracy and the decision of whether or not to delegate authority to bureaucrats in the first place.

Capture

A particular concern of models of the interaction between special interests and the bureaucracy is whether bureaucrats—most particularly regulators—will collude with the citizens or the industry they are meant to regulate. This collusion is commonly termed *capture*. In models of bureaucratic politics, capture is typically modeled using the assumption that bureaucrats have primitive economic preferences and/or are influenced by career concerns. This assumption ensures that bureaucrats may be motivated by bribes or the promise of future employment in the regulated sector.

When models of capture incorporate the agency relationship between the bureaucracy and government, they assume that the government is faced with the problem of moral hazard. If the government had perfect information regarding the actions of the bureaucracy, it would be able to observe, capture, and punish the offenders directly.

Capture may be deterred by legal institutions that investigate the bureaucracy. All else being equal, bureaucrats will be less willing to accept bribes as the risk of detection and punishment rises. Capture may also be deterred by the payment of an *efficiency wage*—a wage above what bureaucrats would receive from alternative employment. The efficiency wage raises the costs to the bureaucrat of being removed from office and thus deters malfeasance.

Appeals to Overseers

Special interests need not resort to bribery to influence the bureaucracy. They may be able to influence bureaucratic decision making simply through the use—or threatened use—of oversight mechanisms. The threat of appeals to the government or to the courts may be sufficient to dissuade the bureaucracy from performing its anointed tasks.

It has long been recognized that special interests may play a role in overseeing bureaucratic performance. Interested members of the public may review bureaucratic performance and report malfeasance or slacking to the government. This form of oversight is frequently termed *fire alarm* oversight and has the advantage of being costless to the government.

However, private sectors entities may be self-interested in their decisions to sound fire alarms. A fire alarm oversight may therefore result in biases in favor of special interests. For instance, an interest group may only report bureaucratic malfeasance if bureaucrats act in a manner that harms that group's interests. Bureaucrats may thus be tempted to favor special interests to avoid review. A government interested in preventing interest-group influence may therefore need to rely on active investigations into bureaucratic decisions in place of, or in addition to, fire alarm oversight. Such investigations are termed *police patrol* oversight.

It is possible that special interests may exercise influence without resorting to the oversight process. The mere threat of appeals to the legislature or courts may prove sufficient to deter regulators. One model of such deterrence assumes that an interest group can signal its willingness to appeal bureaucratic decisions by making campaign contributions to legislators. These contributions act as costly—and informative—signals of the interest group's willingness to contest bureaucratic decisions. The interest group is thus less likely to be regulated as it contributes more to political campaigns.

James Hollyer
New York University
New York City, New York, United States

See also Game Theory; Politicization of Bureaucracy; Principal–Agent Theory; Rational Choice

Further Readings

Epstein, D., & O'Halloran, S. (1994). Administrative procedures, information, and agency discretion. *American Journal of Political Science, 38*(3), 697–722.

Gailmard, S., & Patty, J. W. (2007). Slackers and zealots: Civil service, policy discretion, and bureaucratic expertise. *American Journal of Political Science, 51*(4), 873–889.

Gordon, S. C., & Hafer, C. (2005). Flexing muscle: Corporate political expenditures as signals to the bureaucracy. *American Political Science Review, 99*(2), 245–261.

Laffont, J. J., & Tirole, J. (1993). *A theory of incentives in procurement and regulation.* Cambridge: MIT Press.

McCubbins, M. D., & Schwartz, T. (1984). Congressional oversight overlooked: Police patrols versus fire alarms. *American Journal of Political Science, 28*(1), 165–179.

Moe, T. M. (1989). The politics of bureaucratic structure. In J. E. Chubb & P. E. Peterson (Eds.), *Can the government govern?* (pp. 267–329). Washington, DC: Brookings Institution Press.

Niskanen, W. (1971). *Bureaucracy & representative government*. Chicago: Aldine-Atherton.

Ting, M. (2003). A strategic theory of bureaucratic redundancy. *American Journal of Political Science, 47*(2), 274–292.

Bureaucracy, Street-Level

Throughout the world, legislators and government administrators have increasingly called for greater accountability and improved performance in the delivery of public services, reflecting in part the continuing influence of the theory of new public management that stresses more market-based approaches to service delivery and increased responsiveness to citizens and communities. The worldwide financial crisis has further intensified the pressure on government to improve the efficiency and effectiveness of services.

The performance and accountability movement puts the spotlight on the central role of street-level bureaucrats in the development and implementation of public services. As defined by Michael Lipsky in 1980, street-level bureaucrats comprise workers who interact directly with citizens by providing a valued public service; they include teachers, police officers, case workers, physicians, mental health counselors, and home health workers, and they typically possess a substantial degree of discretion over their work. Importantly though, this discretion is bounded by the restrictions on individual choice placed on street-level bureaucrats by factors common in public and nonprofit agencies, including resource constraints, difficult working conditions, and vague and sometimes conflicting expectations on performance. To cope with these work-related challenges, street-level bureaucrats can sometimes reach decisions that are at variance with the rules and procedures of the agency. Thus, the discretion of street-level bureaucrats can result in policy implementation at odds with the goals and priorities of top administrators, legislators, and the citizenry.

Historically, street-level workers occupied key positions in public agencies. Indeed, the growth of the welfare state in advanced industrial countries in the 20th century typically entailed substantial expansion in the ranks of public sector workers, especially in regions such as the United Kingdom (UK) and the Scandinavian countries. However, in some countries such as Germany and the Netherlands, governments worked in close alliance with large nonprofit organizations who provided social and health services, staffed by countless numbers of street-level workers. The ranks of street-levels workers rose sharply around the world in the 1960s and 1970s with the expansion of the welfare state in policy areas such as health care, social services, criminal justice, and education.

Many of these programs in the United States and abroad encountered serious implementation problems prompting a subsequent reassessment of public services and the role of street-level bureaucrats in service provision. As a result, street-level practice has changed profoundly in the past few decades. First, street-level bureaucrats in social and health care are much more likely to work in a nonprofit or for-profit service agency due to a sharp rise in government contracting for services with these agencies. Indeed, the expansion of services such as home care, community care for the disabled and elderly, substance abuse treatment, and workforce development has been almost entirely through government contracting with private nonprofit and for-profit agencies. While this increase in contracting is particularly pronounced in the United States, it is also evident in many other countries, including the UK, Australia, New Zealand, and Germany.

The growth of contracting has been accompanied by new performance management strategies, including performance contracting, intensified testing, more rigorous practice standards, and outcome evaluation. Overall, these strategies are often an attempt to control the discretion and autonomy of street-level workers. For instance, performance contracts in welfare-to-work programs include specific expectations for street-level workers on the target program goals and sanctions for the agency if these goals are not attained. Similarly, policies promoting "evidence-based practice" seek to enhance the extent to which street-level workers such as mental health counselors or child welfare case workers adhere to professionally recognized standards of best practice. Recurrent fiscal crises of government have also forced public and private agencies to implement new policies to increase the productivity of street-level workers, with corresponding limitations on their ability to depart from established practice standards.

Importantly, the emphasis on productivity by street-level workers is also a ripple effect of intensified competition among service agencies. The paradigmatic street-level bureaucrat of earlier eras possessed autonomy and discretion and was largely immune from market pressures; teachers, child welfare workers, and physicians certainly fit this model well. Currently, though, private nonprofit and for-profit agencies in many countries and jurisdictions actively compete with each other for government grants and contracts as well as private fee income (and in the case of nonprofit agencies, charitable donations). The degree of competition, though, varies tremendously. Rural areas, for example, may lack a sufficient supply of providers to create a competitive market. Government rules and regulations also may affect the supply of providers. Nonetheless, the shift toward more competition in services is a worldwide trend that is evident even in countries with comprehensive welfare state policies and long-standing traditions of private agency and professional autonomy.

Competition in services is also fostered by the growing trend for more individual choice and control in service delivery. The classic street-level bureaucrat working in a large public or nonprofit agency existed within a top-down hierarchical authority system where individual clients had little control or choice in the street-level workers or agencies from whom they received service. In this role, street-level bureaucrats played a key role in constructing the identity of service recipients as well as their access to services and their very understanding of their service rights and expectations.

Changes in the Bureaucrat–Client Relationship

Two developments are changing this street-level bureaucrat–client relationship, although national, regional, and local practices vary tremendously. First, governments are employing a more diverse set of policy tools to address public problems, including vouchers and quasi vouchers and tax credits and deductions. Vouchers in child care, education, housing, and substance abuse treatment can be a more empowering way to engage clients in key decisions regarding their public services. For instance, in community care services for the disabled, individual client accounts have been established by government, allowing the disabled individual to purchase needed services directly, rather than relying on street-level workers. The widespread use of vouchers and quasi vouchers and different forms of payments to private agencies has restructured the contracting relationship so governments have shifted away from direct contracts with an agency to funding support that is tied to the individual citizen users. This funding relationship is inherently more unstable from the contract agency's perspective and puts further competitive pressure on the private service agency and its frontline workers.

The diversification of policy tools is also indicative of a second trend: broad interest and support for individual choice and control in service decisions, evident in vouchers; tax credits for services such as child care; government-subsidized client accounts to purchase needed services; and regulations requiring or promoting greater choice. Over time, this shift is likely to transform the work of street-level bureaucrats who will need to be responsive to client goals and priorities in order to compete effectively for client service dollars.

Greater client choice reflects growing worldwide interest in "coproduction," or the idea that important services such as mental health or community care are coproduced between the street-level worker and individual clients; thus, street-level workers need to pay heed to client motivation and goals in order to achieve successful programmatic outcomes (John Alford, 2009). Coproduction is also related to the concept of "backward mapping," advanced by Richard Elmore (1979/1980), who argued that policymakers should begin the formulation of policy by examining the interactions between clients and street-level bureaucrats. With this information, policymakers should be able to design more effective programs.

In recent years, the coproduction approach has broadened to incorporate more active engagement of clients and community organizations in the work of street-level bureaucrats. Indeed, street-level bureaucrats may need to work with several different public and private organizations on behalf of an individual client. Many service programs are designed at their inception as public-private partnerships with hybrid governance structures representing an array of different local interests and organizations. Thus, the governance of street-level work is vastly more complicated today than in the past, with street-level bureaucrats often engaged with multiple internal and external stakeholders in vertical and horizontal relationships. Consequently, sound policy design is even more important to successful policy implementation today than in the past.

This evolving role of street-level bureaucrats has profound implications for citizenship. Teachers, police, case workers, and emergency shelter staff, to name a few, provide valuable public services that may determine the life chances of individual citizens. As the welfare state grew, so did the ranks of street-level bureaucrats and their importance in providing social benefits such as health care and education essential to full citizenship. Yet street-level bureaucrats can also directly affect the political participation and civic engagement of citizens, so street-level workers also can significantly shape political rights as well.

The capacity of street-level bureaucrats to engage in effective coproduction—and effective services in general—may be severely limited by the intensifying market and accountability pressures facing public and private service agencies. Quality street-level practice requires an ongoing investment by government in training and education. To the extent that street-level workers are engaged in contracting and partnerships, new skill development is often essential. Moreover, the investment obligations of government are not limited to support for public sector workers; the staff and volunteers of private nonprofit and for-profit agencies receiving government contracts require ongoing technical assistance and capacity building to provide responsive, effective services. In an era of scarce resources for public services, street-level bureaucrats will need to work actively with allies in the public and private sectors to create advocates for the value and key importance of public services for all citizens.

Steven Rathgeb Smith
Georgetown University
Washington, D.C., United States

See also Accountability; Bureaucracy; Discretion; New Public Management; Organization Theory

Further Readings

Alford, J. (2009). *Engaging public sector clients: From service delivery to co-production.* London: Palgrave Macmillan.

Elmore, R. (1979–1980). Backward mapping: Implementation research and policy decisions. *Political Science Quarterly, 94*(4), 601–616.

Hill, H. C. (2003). Understanding implementation: Street-level bureaucrats' resources for reform. *Journal of Public Administration Research and Theory, 13*(3), 265–282.

Hupe, P., & Hill, M. (2007). Street-level bureaucracy and public accountability. *Public Administration, 85*(2), 279–299.

Lipsky, M. (1980). *Street-level bureaucracy*. New York: Russell Sage.

Maynard-Moody, S., & Musheno, M. (2003). *Cops, teachers, counselors: Stories from the front lines of public service.* Ann Arbor: University of Michigan Press.

C

Cabinets

Cabinets are small organizational units within governments that consist of politicians who are responsible for the overall policy performance of the government. This entry first describes the history of cabinets and their functions under various forms of government. It then discusses a framework for the analysis of cabinets, which must consider the structural characteristics of the political system, the roles of political parties and of individual ministers, the behavior of the prime minister, and the power of the prime minister's office. Last, the entry describes important tasks for future research on cabinets worldwide.

Historical Origin and Functions

The notion of a cabinet derives from the designation "Cabinet Council," first given in Britain by the king in 1622. The monarch counted on this inner core of privy counselors to legitimize his or her political decisions and thereby his or her duration in office. With the emergence of the modern state in the 19th century, monarchs lost power, and newly developed political parties subsequently replaced their authority. As stressed by John Mackintosh (1968), after these mass parties had come into existence in the 1860s and 1870s, the contemporary structure of the modern cabinet system became visible. Today, a cabinet is a small group of politicians (mostly ministers) responsible for political decision making in countries that have a form of government known as cabinet government (i.e., in Western and Central Eastern Europe and much of the Balkans; in many Commonwealth countries such as Canada, Australia, New Zealand, India, Malaysia, and Singapore; in most ex-British Caribbean and Pacific Islands; in Japan; and in Israel).

Cabinets may also exist in constitutional presidential systems (typically in the Americas), in authoritarian presidential systems (e.g., in Africa and in parts of Asia, including the Middle East), or in traditional monarchical systems (e.g., in the Middle East). In these political systems, however, members of the cabinet serve purely as advisors to the president or the monarch. They are not—as in cabinet government systems—dependent on the majority support of parliament for their duration in office or for getting legislative proposals approved.

In cabinet governments, individual members hold a collective responsibility, which stipulates that all cabinet members are bound to cabinet decisions. They may disagree with these decisions in private but must agree in public. As chairman of the cabinet, the prime minister wields a power that is generally seen as superior to the other cabinet members ("first among equals"). The cabinet members, at least its chair, are responsible to parliament and are expected to control the executive and the legislative branches.

Such cabinets can function satisfactorily only if an equilibrium is found between two opposite requirements—those of representativeness and of efficiency. As Jeffrey Cohen (1988) notes, the requirement of representativeness means that decision making is in the hands of the whole cabinet and not, as, for instance, in the presidential system, in the hands of the leader of the government alone. A cabinet government has therefore to be "collective": Decisions are taken "together." This requirement of "togetherness" is regarded as central to the cabinet system because such a mode of operation is felt to be a more liberal, more democratic, and therefore a superior form of decision making.

Yet this representative-cum-collective principle runs directly against the requirement of efficiency, as efficiency entails that decisions be taken speedily despite the continuous increase in the number and complexity of these decisions. Since representativeness entails togetherness, decisions are likely to be taken only after substantial discussions. This may mean delays in elaborating compromises or even, in cases of extreme conflict, the fall of the cabinet. Bernard Grofman and Peter van Roozendaal (1997) note that because the fall of a cabinet is mostly associated with political crises, cabinet stability becomes the key test of the success of all cabinet systems and the number-one indicator of the performance of cabinets.

Framework for Analysis

In the contemporary literature, such as the work of Jean Blondel and Ferdinand Müller-Rommel (1997, 2001), cabinets are analyzed in the context of the constitutional setting, of the cabinet structure, of the external and internal life of cabinets, and of the decision making among their members.

Constitutional Setting

The general setting of cabinets is provided in broad terms by constitutions. Although constitutions usually give only a little attention to the organization of cabinets, they typically devote more space to matters of appointment and dismissal of cabinet ministers, in particular to the role of the head of state in this respect. Constitutions may, for instance, give the head of state the power to participate in the formation of cabinets (as in France, Finland, Belgium, Bulgaria, and the Baltic States and in Poland until 1993). Constitutions may also foster cabinet stability, for instance, by making it relatively difficult for parliaments to pass "no confidence" motions against the prime minister or other ministers. Future research needs to analyze more systematically the constitutional powers of presidents, prime ministers, and ministers in different cabinet systems.

Internal Operation of Cabinets

The cabinet structure is foremost determined by its size and by its organizational subunits. Most cabinet governments tend to have around 20 members on average. Because these cabinets are too large to be fully involved in all aspects of decision making, cabinet meetings must be well structured and well prepared. In this context, prime ministers, cabinet committees, and cabinet secretariats play an important role. Although the power of prime ministers varies significantly across countries, they are usually more influential in cabinets based on a majority party than in coalition cabinets. Cabinet committees have been set up in many cabinets to speed up the decision-making process. These committees are different in size and status, but they usually include relevant ministers, junior ministers, and civil servants who prepare filtered suggestions for cabinet decision. Some committees are composed only of civil servants who aim to iron out problems that might occur among the different governmental departments. Cabinet secretariats or prime minister's offices, on the other hand, ensure efficiency in cabinet decision making. They control the flow of proposals coming to and emanating from cabinet meetings and resolve problems over proposed agenda items in advance of the cabinet meeting. Studies conducted in parliamentary democracies have shown that the impact of prime ministerial staffs on cabinet government can be excessive. The functions of these staffs range between an administrative and a political pole. These officials act administratively when organizing meetings of the cabinet and the flow of business between prime minister and ministers. They also play a part in the development of long-term ideas about cabinet activities and on the improvement of cabinet decision making, but these forms of activities are never wholly administrative. In most cases, they affect the political operation of the cabinet. Future research must therefore consider, in a systematic and comparative manner, not just personnel and budget but also the actual tasks of the prime ministers' offices.

Impact of External Actors on Cabinets

The operation of cabinets is strongly affected by political parties, parliaments, bureaucracy, and other external actors. Political parties clearly have an impact on the formation, the composition, and the development of cabinets. This impact depends on the overall number of parties in the cabinet (single-party vs. multiparty government) and the type of cabinet coalition (e.g., with or without a dominant party, minority vs. majority party governments). There is usually a stronger battle between cabinets and parties in multiparty coalitions than in one-party majority governments.

The impact of parliament, civil servants, and other political groups on cabinets is closely tied to the role of parties and to the character of coalitions. In some countries, parliaments have substantial power; for example, they may be able to prevent political parties from revolting or erupting. This appears to be particularly true in coalition governments. The impact of civil servants, on the other hand, depends on their position within the administration. The higher civil servants are ranked, and the closer they are to the ministers, the greater their impact on the preparation of cabinet decisions. The role of other external groups in the life of the cabinet can be seen as similar to that of political parties. The most important interest groups are usually linked to specific parties and thereby exercise an indirect impact on cabinet proposals.

Decision Making

Cabinet decision making takes place in the context of a conflict among many patterns of relationship. There are, for instance, differences in the extent to which cabinet meetings are genuinely involved in decision making rather than simply ratifying proposals made by other bodies. The efficiency and effectiveness of cabinet decision making to a large extent depend on the ministerial experience and behavior of prime ministers. Studies on cabinet government have shown that ministers with greater political experience or greater expert knowledge can be expected to exercise more influence on cabinet decision making than other cabinet members. Future research on cabinets has therefore to collect more information on the social background and the careers of ministers under comparative perspective. Further, to understand ministerial behavior, one must know what ministers think about their own work, the work of their colleagues, and the part played by the cabinet meeting itself. If cabinet decisions matter, cabinet members also matter; it is therefore essential to find out how ministers view their role in cabinet government. This means interviewing (former) ministers—an enterprise that is long and complex but is surely rewarding. Cabinet decisions are also affected by the behavior of prime ministers as heads of cabinets. Prime ministers are "arbitrators" or "activists" with respect to the implementation of cabinet policy goals. If they are arbitrators, prime ministers may not be involved personally in policy proposals and must be content with moving the cabinet toward an acceptable solution. If they are activists, they are likely to be deeply interested in particular proposals and to attempt to push these proposals through cabinet. A systematic comparative investigation of cabinets has therefore to develop a typology of prime ministerial leadership style and to relate this typology to the different phases and aspects of cabinet decision making.

Directions for Future Research

Given the difficulties that most cabinets in democratic countries currently face, a common theory about the functioning of cabinets needs to be developed. This theory needs to take into account five key factors that are responsible for the general functioning of cabinets in democratic countries: (1) the structural characteristics of the political system, (2) the role of the political parties, (3) the role of individual ministers, (4) the leadership style of the prime minister, and (5) the administrative characteristics of the prime minister's offices. Systematic research on the functioning of cabinets has first to focus on these five aspects in succession to obtain reliable and comparable information about cabinet governments worldwide. Meanwhile, hypotheses linking these five aspects need to be formulated and tested. This is a vast undertaking, especially since the principles of classification still have to be elaborated in a number of ways. If such an approach is adopted, however, it will become possible to discover and assess the explanations that account for the operation of cabinets.

Ferdinand Müller-Rommel
Leuphana Universität Lüneburg
Lüneburg, Germany

See also Bureaucracy; Government; Parliamentary Systems; Presidentialism

Further Readings

Blondel, J., & Müller-Rommel, F. (Eds.). (1997). *Cabinets in Western Europe.* Houndmills, UK: Macmillan.

Blondel, J., & Müller-Rommel, F. (Eds.). (2001). *Cabinets in Eastern Europe.* Houndmills, UK: Palgrave.

Cohen, J. E. (1988). *The politics of the U.S. cabinet.* Pittsburgh, PA: University of Pittsburgh Press.

Grofman, B., & van Roozendaal, P. (1997). Modelling cabinet durability and termination. *British Journal of Political Science, 27*(3), 419–451.

Mackintosh, J. J. (1977). *The British cabinet* (3rd ed.). London: Stevens.

CAPITALISM

Since the early 20th century, the concept of capitalism has been used to capture the structure and dynamics of a particular historical formation of economy and society that first emerged during the late Middle Ages in Southern Europe and later spread to Northwestern Europe. Capitalism emphasizes the attainment of profit through the operation of the market and private ownership of the means of production. This entry describes the role of markets and property rights in capitalism and examines the functions of firms and institutions in its operation. The characteristics of capitalist culture are then discussed. The entry closes with a review of endogenous critiques of capitalism and of the shape it may take in the future.

Since its beginnings in Europe, capitalism has expanded to virtually all parts of the globe. Capitalism can be contrasted with subsistence economy, feudalism, socialism, and slave economy. Third World developing societies may contain insular capitalist patterns in their economy without thereby becoming capitalist *societies*. Comparative social scientists and historians have distinguished a great number of stages, types, qualifiers, and variants under the broad umbrella concept of capitalism, such as agrarian, commercial, industrial, and financial capitalism; state capitalism and coordinated capitalism; and Nordic, Anglo-Saxon (chiefly British and American), East Asian, and "Rhenish" capitalism.

Most scholars who use the concept of capitalism in a holistic way view it not only as an economic system but also as comprising a type of social structure, political institutions, and specific cultural norms and values. The complementarity, goodness of fit, and range of variation that exist among these realms—essentially the realms of capitalist interests, institutions, and ideas that together make up capitalism—have been the focus of social science analysis since the pioneering works, around the turn of the 19th century, by Max Weber and Werner Sombart to contemporary research on comparative capitalism. On the European continent, the use of the term *capitalism*, in both political and academic contexts, almost always had critical overtones. Authors who wish to avoid such connotations use "social market economy," "industrial society," or simply "modern society" instead; however, such usage may cause them occasionally to lose sight of the *problèmatiques* and insights of those classical authors of social science.

Defining Features of Capitalism

There are six defining features of capitalism: (1) *markets*, (2) *property rights*, (3) the role of private *firms*, (4) politico-economic *institutions*, (5) capitalist patterns of the cognitive and normative *culture* (Weber's "spirit" of capitalism), and (6) reflexive dynamics of *critique* that are specific to capitalist societies. Theorists differ as to the emphasis they attach to each of these components of capitalism. The study of capitalism is a highly interdisciplinary field of investigation to which historians, economists, sociologists, lawyers, political scientists, and philosophers have significantly contributed.

Markets

Capitalist societies are based on economic systems in which most goods and services are bought and sold in markets for a monetary price, thus making them commodities. The commodification of goods makes for the contingency of economic transactions, meaning that the parameters of these transactions—who buys from and sells to whom and what commodity at which point in time and at what price—become a matter of continuous choice and an ongoing competitive recombination of social relations. Market transactions are governed by a regime of social norms and legal rules (law of contract) that is enforced by a neutral state-operated court system and that specifies the mutual rights and liabilities of agents entering into economic transactions. These norms and rules are designed to rule out the use of openly predatory practices such as individual or organized violence, fraud, theft, deception, conspiracy to deny potential suppliers market access, to some extent even practices of cartelization and monopolization, and bribes as "unfair" means employed in the pursuit of economic gain. To the (historically highly unlikely) extent that such rules are fully implemented, we can speak of a civilizing function of market competition, ideally leaving only prices and qualities/novelty of goods as action parameters of competitors. Yet as prices are "given" in any (nearly) perfect (or "atomistic") market, suppliers are under strong incentives to innovate both products and (technical and organizational) production processes. Markets determine prices in response to changes in the volume of supply of and demand for specific goods and services. The prominent role of choice, contingency, and civility in economic interaction has led theorists such as Milton Friedman (1962) to equate capitalism with individual freedom. Markets make people "free to choose."

Yet a defining feature of capitalist market societies consists in the fact that not only goods and services manufactured for the purpose of being sold but also the "factors of production" employed in the process of manufacturing are subject to market exchange under capitalism. Even though these factors—natural resources, human labor power, and money—have not been "produced" (and certainly not produced with the intention of being marketed), they are still subsumed under the commodity form. This "commodification of noncommodities" has been attacked, both in the Marxist tradition and famously by Karl Polanyi (1944), as the core contradiction of capitalism—that is, a source of conflict and instability that constantly calls for (interventionist, reformist, revolutionary, authoritarian, military, etc.) remedies and institutional safeguards to be installed by holders of political power. The commodification of money through speculative investments, with their potentially disastrous effects on financial markets; the commodification of natural resources with their associated environmental damages; and the commodification of human labor power with its distributional and other adverse impacts can all be cited as contemporary instances of this key "mistake" of commodifying what by their very nature are noncommodities, or "fictitious" commodities. Thus, one defining feature of a capitalist society and its dynamics is the existence of a labor market in which the capacity of workers to perform productive services is being traded under labor contracts.

Such commodification of noncommodities has provided for the enormous gains in efficiency, growth, and prosperity that have accompanied the history of capitalism. While under capitalism the commodity form is extended to noncommodities, it is on the other hand restricted (compared with precapitalist monetized exchange relations) to items of "economic" value—that is, excluding items such as positions in the state administration, court decisions, academic titles, marriage licenses, or, most important and since the abolition of slavery, human beings themselves, who are instead governed by the principle of inalienable "self-ownership," a concept that has its roots in the work of theorists such as John Locke.

By exposing labor to market contingency under the regulatory regime of the labor contract, capitalism inserts workers into the organizational framework of productive organizations (firms), which, due to the division of labor, organizational control mechanisms, and efficiency-enhancing investment goods, allow labor power to be used much more productively than was the case in precapitalist forms of production. Yet the reverse side of this growth and prosperity lies in the commodification of (nominally "free") contractual labor and the distributional patterns as well as the contingencies following from it. Not only does the individual employer exercise power (as authorized by the labor contract) over the employees within the firm, but the collectivity of all employers also exercises "structural" or class power over the labor force as a whole. The latter arises from the fact that labor is tied to capital in a relationship of asymmetrical dependency. Because wages are normally workers' only source of subsistence, workers typically depend more urgently on being employed than employers depend on employing labor, or at least on employing it "here, now, and continuously." One important cause of this asymmetry resides in the fact that employers are in control of (as well as incentivized through competition to put to use) productivity-enhancing and hence of labor-saving technical and locational change, whereas workers, by themselves, can do little (if anything) to enhance the "welfare yield" of the wages they earn and spend. The process that has been set up by the fundamental capitalist institutions of the labor market and the labor contract can thus be looked on as a "wealth-maximizing game" and, at the same time, a "poverty- and insecurity-generating game."

After the end of the golden age of stable growth and full employment—an age that coincided roughly with the third quarter of the 20th century for the nations in the Organisation for Economic Co-operation and Development (OECD)—one of the core problems of open (globalized), advanced capitalist economies has become the chronic imbalance between the supply of labor and the demand for it, as was pointed out by Guy Standing (2009). This core problem translates into the divisive dilemma often confronting workers, who must either yield "flexibly" to the pressures to become—and remain—"employable" (which they can rarely accomplish through their own means and efforts alone) or face the prospect of socioeconomic insecurity, precariousness, unemployment, and exclusion.

Property Rights

Participants in markets enjoy state-enforced property rights. More important, such property rights do not imply that every member of the legal community has a right to own some share in the total of the material resources available in the community (as was envisaged by 18th-century revolutionary writers such as Thomas Paine). Rather, it means that members of the community who already happen to be in the possession of units of property are recognized and protected in their ownership status (provided that such property has been acquired in legally permitted ways). Such protection means that the ownership status is safeguarded against the loss of property (e.g., through theft, destruction, confiscation, etc., although, of course, not against losses that may result from unfavorable choices made by the owners). The right to property further means that the owner is free (within limits established by regulatory law) to determine the use to which the property is being put as capital, as well as to appropriate the gains (profits) flowing from its use. One highly consequential aspect of the freedom of owners as constituted by property rights is to use these rights to hire labor, with the further implication that those hired as workers are rendered incapable of acquiring property themselves, under the terms of remuneration under which they are hired and in the absence of potentially productive property of their own. In this case, the availability of property rights to some does not simply coexist with but causes the denial of property rights to others (particularly as the latter typically lack the collateral that they would need to obtain a commercial bank loan).

Firms

If property owners decide to invest their monetary resources (which is the only chance they have to make their property a durable asset or even increase it through accumulation), this investment will show up as capital in firms, unless it is a purely speculative investment in financial markets. The firm is the key capitalist institutional location where investment (in buildings, machinery, raw materials, etc.) is combined with wage labor for the purpose of producing marketable commodities. The two defining features of the capitalist firm are (1) its distantiation (in space, time, social function, and accounting principles) from the household and (2) a hierarchical structure of command and control (with an entrepreneur or a managing board at the top) and a formalized vertical and horizontal division of labor designed to promote the efficiency of the productive process and the realization of its results in markets. In sharp contrast to voluntary and highly contingent market transactions, the interaction (which Marx called "despotic") that occurs within firms is based on the contractual right of some agents to give orders and the duty of others (as enforced by some measure of the right of superiors to terminate the labor contract) to carry them out. While the degree of authoritarianism of the internal regimes of capitalist firms may vary greatly, some measure of authoritative coordination seems indispensible. This is so because, first, the labor contract is (for good reasons in terms of efficiency) essentially an incomplete contract with gaps that need to be filled by commands and, second, the conflict of interest inherent in any asymmetrical contractual relation will make it highly unlikely that those gaps in the contract will reliably be filled by acts of spontaneous cooperation, thus overcoming this tension. The result is friction between the (nominally) voluntaristic, egalitarian, and freely chosen relationship between firms and their external market partners, on the one hand, and the authoritarian and hierarchical relations inside the firm as a formal organization, on the other.

The existence of this relationship of social power can be factually accounted for in terms of the asymmetry that capital can hire labor; yet labor, in the absence of savings or a collateral, cannot normally hire capital. And it can institutionally account for the fact that, once employees enter into the labor contract, they subject themselves to the command structure of the enterprise. Virtually all of the labor market, labor relations, and industrial relations policies that we find in the history of capitalist societies (including the building of institutions for collective bargaining and codetermination) can be seen as ongoing attempts of states, workers, and employers to regulate and (un)balance these two basic kinds of power relationships. They are analytically distinct from a third one: the political "exit power" of capital to relocate (or to threaten relocation or shifts to speculative investments in financial markets), to which states and their political elites are vulnerable. This vulnerability results from the dependence of states on tax revenues obtained, directly or indirectly, from investment, growth, and employment, all of which are ultimately controlled by decisions of capitalist enterprises.

The aggregate effect of activity in firms and labor markets generates and reproduces specifically capitalist patterns of inequality, reflecting the differential marketability (employability) of labor. As Standing (2009) notes, these patterns pertain to earnings, employment opportunities, income security, wealth, organizational resources, political power, and even life expectancy. These inequalities unfold in interindividual, intersectoral, interregional, international, and global dimensions. At the bottom of distributional hierarchies are people, regions, and even an entire continent (Africa) whose survival is placed in jeopardy by the dynamic of capitalism or whom the dynamic of capitalism can afford, as it were, "to do without."

Firms are the institutional location where a specific kind of capitalist rationality unfolds. At every point in time, the intellectual technique of rational capital accounting of costs and returns allows a firm to assess its own value and to evaluate alternative courses of action in terms of probable returns to the capital invested. At the same time, there are two sources of irrationality, one at the source and one at the outcome of rational calculation. As to the former, accumulation is seen as originating from a deeply irrational dynamic of entrepreneurial intuition and vision that can be neither taught nor learned—a kind of imaginative "creativity at the top" that guides the innovation of products and processes. As to the latter, the aggregate outcome of the capitalist dynamic triggers social (including environmental and cultural) changes that just "happen" as unintended and unpremeditated outcomes and cannot be attributed to any rational design or calculation. Even firms themselves can never be certain, given the unpredictability of their environment, that decisions will turn out to have been rational in light of the outcomes. This latter irrationality of outcomes is a point at which the dynamics of capitalism and the normative ideals of modernization diverge. If the project of modernity means the effective mastery of society over its collective fate, this is clearly not a virtue in which capitalism (according to Max Weber, arguably still the "most fateful force of social life") excels.

Institutions

Firms interact not only with external market participants (customers, workers to be hired, and other firms as suppliers or buyers) but also, and in ways that are not mediated by markets, with the institutional environment in which they are "embedded." Sociologists and "institutionalist" economists have explored the vast field of noncommercial interaction that both firms as employers as well as workers as employees are involved in. These noncommercial relations of capitalist firms, as well as of everyone else participating in markets, are governed by institutions and the legal rules, formal procedures, and social norms that institutions consist of. The institutional environment of market actors (beginning with private law and its enforcement in courts and not ending with state-provided investment in infrastructure) is something that they, on the one hand, depend on to reach their market objectives; yet on the other hand, the institutional environment is one that market actors try to actively shape and transform in ways that best suit their respective interests. They are involved in a reflexive process that may be termed the *production* of the conditions of production (as well as their *distribution*). The politically mediated production of conditions of production applies to the areas of research, development, and technical change, for instance, in the areas of communication, transportation, and the development of new materials and sources of energy. It also applies to the vast policy areas of infrastructure investment, regulation of markets, trade policy, taxation, labor market and social policy, and macroeconomic steering. At any rate, we would get a seriously deficient and distorted picture if we were to model action in capitalist societies as primarily the market action of buying, selling, and investing. To succeed, market actors take an equally strong interest not only in complying with but also in strategically shaping the noncommodified environment of commercial interaction in markets. The relationship between the actors of a capitalist economy and the institutional environment in which they act is a reciprocal one: Firms, consumers, owners, and workers would not be able to make a single move without relying on premises such as laws, courts, legislatures, regulatory agencies, police protection, schools, physical infrastructure, systems of taxation and tariffs, social insurance systems, central banks, research and development organizations, and many others, mostly supplied, sponsored, and regulated by state agencies and all kinds of private–public hybrids, which in recent literature are referred to as agencies of "governance." Although capitalism is a global system, the configuration of capitalist economic actors and their institutional environment appears to be still largely shaped by national traditions, path dependencies, and policy approaches shared by the political and economic elites of nation-states. These institutional-context conditions are never "given," fixed, or sacrosanct according to a master formula of the "mixed economy" but are, in fact, in constant flux under the impact of hegemonic doctrines of social order (e.g., neoliberalism) as well as the strategic efforts of economic actors to alter them in ways that allow for better exploitation of the emerging opportunities. These agents depend on an institutional framework of social order, yet at the same time they are constantly involved in strategic activities designed to disorganize and reorganize it. The capacity for the latter is derived, in spite of the apparent primacy of political state power over economic exchange, from the fact that modern states, in particular modern, liberal, democratic states, and their stability depend as much on the reasonably smooth operation and growth performance of the capitalist economy as the agents in the capitalist economy depend on the state-provided institutional setup. Again, there is an asymmetrical mutual dependency (in contrast to the notion of a hierarchical primacy of the state over the economy), due to the fact that (capitalist) states, in their turn, depend on both fiscal resources and political support (with labor market outcomes as one of its important determinants) for the sake of their stability. This dependency of the state on capital and its profitable investment is all the greater, and the state's vulnerability more significant, the more the state is a welfare state (i.e., a state with substantial legal commitments to the provision and maintenance of social security) and the more investors enjoy the "exit option" that a denationalized pattern of trade and investment (globalization) provides.

What agents under capitalism actually do is thus much more than buying and selling in the pursuit of gain, profit, and utility maximization. Beyond that, they act reflexively on the very institutional context conditions under which they act, revising, as it were, the rules as the game goes on by constantly reframing what Robert Boyer and Yves Saillard (2002) call their "accumulation regimes." Firms and their associations are involved in the legal or managerial design and ongoing adjustment of "production regimes" and modes of "corporate governance." Moreover, they form cartels and alliances, make threats and promises, warn and demand, bargain and negotiate, associate, advertise, influence, lobby, launch campaigns, donate, resist, mobilize, implicitly blackmail political authority by virtue of the fact that some economic agents are "too big to die," complain about state policies and advocate alternative ones, opportunistically evade legal and contractual obligations, strike political deals, and so on—all to shape, reshape, and occasionally also subvert the institutional context within which the core economic process of capitalism and the competitive pursuit of profits is going on. Capitalism is a political economy in the sense that it can hardly be conceptualized in terms of a durable institutional equilibrium. To the contrary, rules and their recognition are permanently contested. If the state and its institutions can be said to be devices to generate security of expectations leading to trust and to protect capitalist market society from its own inherent dangers of destabilization, it can also be said that this device is itself not reliably protected from the repercussions of such destabilization. Charles Lindblom (1982) has even compared the market to a "prison" in which the makers of public policies are incarcerated. The assumption that capitalism is at all "governable" (as opposed to essentially "anarchic") is, in other words, far from axiomatic. This condition of uncertain institutional embeddedness applies even to the overall political regime type. For while it is true that all liberal democracies contain capitalist economies (in spite of the friction that exists between the two), the reverse is not true: Both historically and in the contemporary world, capitalism has coexisted with (and indeed flourished under) various types of nondemocratic regimes.

Capitalist Culture

What Max Weber has termed the *spirit* of capitalism is a complex and multifaceted phenomenon that includes cognitive and epistemic as well as motivational and justificatory elements. It has in part motivated the transition from precapitalist to capitalist modes of economic life, and it is in turn shaped and inculcated by the realities of life under capitalism. It also applies differently to different types of actors (e.g., manual workers, white-collar workers, the self-employed, entrepreneurs, managers, and consumers) in the capitalist game, as well as to different stages of its development and associated "production regimes." For example, Boyer and Saillard (2002) contrast "Fordist" mass production versus post-Fordist "flexible quality production." Central to the core model of capitalist culture is the notion of selfish and "unfraternal" (as Max Weber put it) individuals' pursuit of acquisitive rationality for its own sake. These individuals methodically explore the physical and social world in constant search of opportunities for gain. In doing so, they follow their interests, control their own passions through self-imposed discipline, and resist the passions of rulers. This pursuit of interest is conceived as endless—both in the sense that there is no end or state of satiation to be reached where further efforts become pointless and in the sense that it can (and in fact must) go on forever, as any standstill spells failure in a competitive environment. The rationality that governs this behavioral dynamic is "formal," "abstract," "self-referential," unendingly and relentlessly expansive in time and space, and boundless. Everything we encounter in the world is first of all being framed in terms of costs and returns, risk of loss, and opportunity for gain alone. At the same time, the accounting frame of capitalist assessment of costs and returns is too restricted and myopic—that is, insufficiently intelligent—to capture long-term and collective negative externalities, which therefore tend to be systematically ignored. Weber has claimed an "elective affinity" that exists between the urge to accumulate and Puritan asceticism, which abhors wealth to be enjoyed and instead lauds its being transformed into capital to be invested, with the satisfaction of need just being a by-product of the process. Weber, in his 1904 book *The Protestant Ethic and the Spirit of Capitalism*, argued that the religious values of Calvinism provided motivation for work and the accumulation of wealth in the secular world. He saw this as an alternative to Marx's doctrine of historical materialism as the basis of economic structures. (Marx almost never used the term *capitalism*.) Relentless and often fear-driven patterns of acquisitive search behavior have become common today not just among entrepreneurs and the self-employed but also among a wide range of employees who have adopted entrepreneurial attitudes and values such as the values of flexibility and mobility.

But this characterization of capitalist culture and cognitive style captures just one of its facets. As Daniel Bell (1976) pointed out, others include patterns of hedonistic consumerism with its mindless shortsightedness that creates ever more needs and desires amid affluence. Still another consists in countertendencies to abstract formal rationality, the spread of which can be observed in the irrationalities of superstition, magical thinking, and prejudice of "authoritarian" modal personalities, or in tendencies in postmodernist and antiauthoritarian countercultures that challenge dominant modes of rationality, thus causing "cultural contradictions of capitalism" (Bell, 1976). Whether or not we can speak of specific cultural prerequisites (as opposed to formative cultural consequences) of capitalism is an issue of considerable interest for social research, given that East Asian capitalism has thrived within the cultural framework of Confucianism and also given that there was hardly any founding generation of a capitalistically "spirited" middle class in some of the now capitalist societies that emerged from state socialism in central East Europe, where capitalism was built "without capitalists," as was stressed by Gil Eyal, Ivan Szelényi, and Eleanor Townsley (1998). As is the case with political institutions, both high culture and popular culture and both the normative and the cognitive "mental infrastructure" of capitalist societies are constantly (re)negotiated through the joint governance of private actors (e.g., much of the media industries) and public authorities.

Critiques of Capitalism

From its historical beginnings, capitalist market societies have encountered strong endogenous critiques. The intellectual and political critique of capitalism and its inherent dynamics comes in two main variants that are often combined by critics. One is based on empirical analysis and prediction and focuses on the observable instability of the systems and its built-in self-destructive tendencies; this kind of critical perspective yields crisis theories according to which the system will become, sooner or later, unsustainable. The other critical perspective is normative and highlights the suffering, deprivation, exclusion, sense of meaninglessness, and various kinds of injustice that are perceived as concomitant features of capitalist growth and development; in response to this experience of injustice, social conflict, be it in the form of class conflict or otherwise, is both predicted and advocated by critics to overcome capitalism and transform it into a type of society that is both more just (at the level of "social integration") as well as more stable as a viable economic system.

However, the empirical observation of cyclical patterns of crisis that unfold under capitalist institutions, as well as the normative focus on injustice, does not provide a robust argument to the effect that we are entitled to anticipate a crisis or a conflictual challenge of capitalism. This non sequitur is widely recognized today as an analytically unwarranted leap of political faith. For just as cyclical crises and recessions set the scene for ever new rounds of accumulation and growth, capitalism can also thrive on at least some versions of its normative critique, thus arguably continually contributing to the system's perpetuation rather than its demise.

Conclusion

It seems safe to state that the notion of a modern society "after" and "without" capitalism and its key features—a notion that has inspired much of the history of the political left—has largely been rendered obsolete today. This obsolescence is epitomized by the demise of European state socialism in 1989 to 1991. Neither the probability nor the desirability of a full-scale historical abolition (breakdown) of capitalist patterns of socioeconomic organization is widely advocated any longer. Instead, capitalism is seen to be subject to numerous forces of endogenous change, leading to a great deal of variation and institutional diversification of capitalisms. The longitudinal notion of a diachronic sequence of types of social order has yielded to a "synchronized" perspective, with post- and anticapitalist, "decommodified," and solidaristic patterns of socioeconomic organization now playing a role in ongoing and contingently reversible modifications of enduring capitalist core structures and the ongoing recombination of its components. As stated before, it is in the nature of capitalism that it consistently breeds reflexive critiques of capitalism. These aim at curbing and holding accountable the various manifestation of the social power of capital and propose to deploy a variety of institutions and policies for its domestication—be it social power over a firm's employees, power at the level of class relations and its distributional consequences, the role of economic (veto) power in the making of public policy, the power of investors to inflict vast negative externalities on everyone else (and even on themselves) through economic crises and environmental destruction, or the power of capital to shape and "colonize" the process of cultural reproduction.

Claus Offe
Hertie School of Governance
Berlin, Germany

See also Critical Theory; Globalization; Individualism; Inequality, Economic; Liberalism; Market Economy; Marx, Karl; Marxism; Neoliberalism

Further Readings

Bell, D. (1976). *The cultural contradictions of capitalism.* New York: Basic Books.
Boltanski, L., & Chiapello, E. (2005). *The new spirit of capitalism.* London: Verso.
Boyer, R., & Saillard, Y. (Eds.). (2002). *Regulation theory: The state of the art.* London: Routledge.
Eyal, G., Szelényi, I., & Townsley, E. R. (1998). *Making capitalism without capitalists: Class formation and elite struggles in post-communist Central Europe.* London: Verso.
Friedman, M. (1962). *Capitalism and freedom.* Chicago: University of Chicago Press.
Hall, P. A., & Soskice, D. W. (Eds.). (2001). *Varieties of capitalism: The institutional foundations of comparative advantage.* Oxford, UK: Oxford University Press.
Lindblom, C. E. (1982). The market as prison. *Journal of Politics, 44*(2), 324–336.
Polanyi, K. (1944). *The great transformation: The political and economic origins of our time.* Boston: Beacon.
Standing, G. (2009). *Work after globalization: Building occupational citizenship.* London: Edward Elgar.
Streeck, W. (2009). *Re-forming capitalism: Institutional change in the German political economy.* Oxford, UK: Oxford University Press.

Case Studies

Although much of what we know about the empirical world has been generated by case studies and case studies continue to constitute a large proportion of work generated by the political science discipline, the case study *method* is poorly understood. Even among its defenders, there is confusion over the virtues and vices of this research design. Practitioners continue to ply their trade but have difficulty articulating what it is they are doing, methodologically speaking. The case study survives in a curious methodological limbo. The problem of ambiguity begins with the term itself. To refer to a work as a *case study* might mean that its method is qualitative, small *N*; that the research is holistic and thick (a more or less comprehensive examination of a phenomenon); that it uses a particular type of evidence (e.g., ethnographic, clinical, nonexperimental, non–survey based, participant observation, process tracing, historical, textual, or field research); that its method of evidence gathering is naturalistic (a "real-life context"); that the research investigates the properties of a single observation; or that the research investigates the properties of a single phenomenon, instance, or example. This entry presents a more precise definition of this method and discusses its strengths and limitations.

Evidently, researchers have many things in mind when they talk about case study research. Confusion is compounded by the existence of a large number of near synonyms—single unit, single subject, single case, $N = 1$, case based, case–control, case history, case method, case record, case work, clinical research, and so forth. As a result of this profusion of terms and meanings, proponents and opponents of the case study marshal a wide range of arguments but do not seem any closer to agreement than when this debate was first broached several decades ago. To talk about this subject in a productive fashion, we must arrive at a narrower definition. In this entry, we stipulate that a *case* connotes a spatially delimited phenomenon (a unit) observed at a single point in time or over some period of time. It comprises the sort of phenomena that an inference attempts to explain. Thus, in a study that attempts to explain certain features of nation-states, cases consist of nation-states (across some temporal frame). In a study that attempts to explain the behavior of individuals, individuals make up the cases. And so forth. Each case may provide a single observation or multiple (within-case) observations.

For students of political science, the archetypal case is the dominant political unit of our time, the nation-state. However, the study of smaller social and political units (regions, cities, villages, communities, social groups, and families) or specific institutions (political parties, interest groups, and businesses) is equally common in other subfields and, perhaps, increasingly so in comparative politics. Whatever the chosen unit, the methodological issues attached to the case study have nothing to do with the size of the individual cases. A case may be created out of any phenomenon so long as it has identifiable boundaries and comprises the primary object of an inference. Note that the spatial boundaries of a case are often more apparent than its temporal boundaries. We know, more or less, where a country begins and ends, even though we may have difficulty explaining *when* a country begins and ends. Yet some temporal boundaries must be assumed. This is particularly important when cases consist of discrete events—crises, revolutions, legislative acts, and so forth—within a single unit. Occasionally, the temporal boundaries of a case are more obvious than its spatial boundaries. This is true when the phenomena under study are eventful but the unit undergoing the event is amorphous. For example, if one is studying terrorist attacks, it may not be clear how the spatial unit of analysis should be understood, but the events themselves may be well bounded. Following this understanding of "case," a case study could be defined as *the intensive study of a single case for the purpose of understanding a larger population of cases*. Several implications flow from this definition, as applied to the social sciences.

By virtue of the small size of the sample and the heterogeneous nature of most social phenomena, the *representativeness* of a case study is always rather uncertain. One does not know for sure whether the chosen case or cases are similar (in relevant ways) to the larger population of theoretical interest. Of course, there is always some uncertainty attached to the quest for external validity. Even so, the uncertainty normal to this sort of inference is compounded whenever the chosen sample is very small. By virtue of the small size of the sample, it can usually be presumed that the evidence at hand is *nonexperimental*. Note that if a treatment can be manipulated, there is usually an opportunity to generate large treatment and control groups, moving the style of research away from an intensive focus on a single case. (That said, it is certainly possible to combine a large-*N* cross-case analysis of experimental evidence with the in-depth study of an individual case, or several cases, within that sample—perhaps with the goal of shedding light on causal mechanisms. Varying approaches to the same evidence often serve a complementary function.)

Case study research may incorporate several cases, as in the style of research known as *comparative historical*. However, at a certain point it will no longer be possible to investigate those cases intensively. At the point where the emphasis of a study shifts from the individual case to a sample of cases, we shall say that a study is a *cross-case* one. Thus, the distinction between a case study and cross-case study must be understood along a continuum. The fewer cases there are, and the more intensively they are studied, the more a work merits the appellation case study. Although the case study has been identified thus far by the smallness of the sample, it is important to appreciate that whenever one is conducting an intensive analysis of a single case (or a small number of cases) one is gathering multiple observations from the case(s). Intensive study means multiple observations, usually at a lower level of analysis. Insofar as these observations are comparable with one another, they may be analyzed in a quantitative fashion. Thus, the case study format is entirely compatible with large-*N* statistical analysis of within-case evidence.

To clarify, what distinguishes a large-*N* analysis embedded within a case study from a large-*N* cross-case analysis is that the former is nested within the units of primary theoretical interest. Thus, if one's purpose is to explain the process of national-level democratization, the units of theoretical interest are nation-states. A case study approach to this question would focus on a single country or a small sample of countries. As part of such a study, one might conduct a survey with thousands of respondents to determine attitudes and behaviors relevant to democratization. This provides the fodder for a quantitative analysis, nested within the case study format. If, on the other hand, one's theoretical purpose is to explain *individual-level* variation in attitudes/behavior related to democratization, then the same study is probably best classified as cross-case, rather than case study: One now has a large sample of the things to be explained, none of which receives a great deal of attention.

So conceived, a case study analysis may contain qualitative or quantitative evidence. Typically, it combines both. So conceived, the case study method is consonant with any theoretical framework and virtually any methodological framework—for example, behavioralism, ethnography, rational choice, institutionalism, interpretivism, and, very occasionally, experimental research ($N = 1$ experiment). So conceived, it is possible to look dispassionately at the strengths and weaknesses of the case study research design, in contrast with cross-case research designs. These trade-offs derive, first of all, from basic research goals such as

1. whether the study is oriented toward hypothesis generating or hypothesis testing,
2. whether internal or external validity is prioritized,
3. whether insight into causal mechanisms or causal effects is more valuable, and
4. whether the scope of the causal inference is deep or broad.

These trade-offs also hinge on the shape of the empirical universe—that is,

5. whether the population of cases under study is heterogeneous or homogeneous,
6. whether the causal relationship of interest is strong or weak,
7. whether useful variation on key parameters within that population is rare or common, and
8. whether available data are concentrated or dispersed.

Along each of these dimensions, case study research usually has an affinity for the first factor and cross-case research has an affinity for the second. All else being equal, case studies are more useful when the strategy of research is exploratory rather than confirmatory/disconfirmatory, when internal validity is given preference over external validity, when insight into causal mechanisms is prioritized over insight into causal effects, when propositional depth is prized over breadth, when the

Table 1 Case Study and Cross-Case Research Designs: Affinities and Trade-Offs

	Affinity	
	Case Study	*Cross-Case Study*
Research goals		
1. Hypothesis	Generating	Testing
2. Validity	Internal	External
3. Causal insight	Mechanisms	Effects
4. Scope of proposition	Deep	Broad
Empirical factors		
5. Population of cases	Heterogeneous	Homogeneous
6. Causal strength	Strong	Weak
7. Useful variation	Rare	Common
8. Data availability	Concentrated	Dispersed

Source: Gerring, J. (2007), p. 38. Copyright Cambridge University Press; used by permission.

population of interest is heterogeneous rather than homogeneous, when causal relationships are strong rather than weak, when useful information about key parameters is available only for a few cases, and when the available data is concentrated rather than dispersed. These trade-offs are summarized in Table 1. To clarify, they represent methodological *affinities*, not invariant laws. Exceptions can be found to each one. Even so, these general tendencies are often noted in case study research and have been reproduced in multiple disciplines and subdisciplines over the course of many decades.

It should also be stressed that each of these trade-offs carries a ceteris paribus caveat. Case studies are more useful for generating new hypotheses, *all other things being equal*. The reader must bear in mind that many additional factors also rightly influence a writer's choice of research design, and they may lean in the other direction. Ceteris is not always paribus. One should not jump to conclusions about the research design appropriate to a given setting without considering the entire range of issues involved—some of which may be more important than others.

In conclusion, it may be granted that case studies and cross-case studies explore the world in different ways. Yet, properly constituted, there is no reason why case study results cannot be synthesized with the results gained from cross-case analysis and vice versa. Indeed, they are often complementary. In this light, one might cite the current popularity of multimethod work (triangulation), which often combines large-*N* cross-case analysis with an intensive focus on one or a few carefully chosen cases. Although this discussion has been brief, it may help restore a greater sense of coherence, purpose, and integrity to the case study method. This narrower and more carefully bounded definition may alleviate some of its most persistent ambiguities. It is also hoped that the characteristic strengths of this method, as well as its limitations, will be apparent to producers and consumers of case study research. The case study is a useful tool for some research objectives but not all.

John Gerring
Boston University
Boston, Massachusetts, United States

See also Analytic Narratives: Applications; Analytic Narratives: The Method; Configurational Comparative Methods; Mixed Methods; Participant Observation; Process Tracing; Quantitative Versus Qualitative Methods

Further Readings

Eckstein, H. (1975). Case studies and theory in political science. In F. I. Greenstein & N. W. Polsby (Eds.), *Handbook of political science: Vol. 7. Political science: Scope and theory* (pp. 94–137). Reading, MA: Addison-Wesley.

George, A. L., & Bennett, A. (2004). *Case studies and theory development.* Cambridge: MIT Press.

Gerring, J. (2007). *Case study research: Principles and practices.* Cambridge, UK: Cambridge University Press.

Categorical Response Data

The term *categorical response data* refers to data for outcome variables whose values represent distinct categories as opposed to continuous quantities. Categorical variables are pervasive in political science. Examples are as diverse as party affiliation ("Democrat," "Republican," "other," "none"), union membership ("yes," "no"), form of government ("republic," "monarchy," "military dictatorship," etc.), voter participation ("yes," "no"), or confidence in the government ("a great deal," "quite a lot," "not very much," "none at all"). Although examples for the collection of categorical data can probably be traced back as far as the ancient censuses in Egypt, it was not until the early 20th century that a systematic development of statistical methods for the analysis of categorical data began. The starting point was the work of Karl Pearson and G. Udny Yule, who debated over how best to analyze associations between categorical variables. Many fundamental contributions that prepared the ground for the emergence of a wide variety of approaches in the subsequent decades were also made by R. A. Fisher in the 1920s and 1930s. This entry provides a brief overview of the rich collection of methods for the analysis of categorical data available today. First, a description of different types of categorical data is given. Second, various ways to analyze categorical data are summarized.

Typology of Categorical Data

The most important differentiation of categorical data distinguishes between *nominal* and *ordinal* variables. The values of a nominal variable identify categories that do not have a natural rank order. Examples are ethnicity ("Black," "White," "Hispanic," etc.) or marital status ("single," "married," "divorced," "widowed"). Methods using nominal data should not depend on the numerical values assigned to the categories. The categories of ordinal data, in contrast, possess a natural ranking. Examples are social class ("lower," "middle," and "upper"), agreement with the statement "Democracies are indecisive" ("strongly agree," "agree," "disagree," and "strongly disagree"), or self-reported political left–right orientation (on a scale from 1 = *left* to 10 = *right*). Methods using ordinal data take into account the order of the values, but the exact spacing does not matter. The distinction between nominal and ordinal variables is not always clear. For example, political parties can be ranked on a left–right scale, but the order may be ambiguous for some parties (because they may rank differently on different policies). Moreover, whether a variable is treated as ordinal can depend on the research question.

A further distinction is made between *qualitative* and *quantitative* data. Some authors use the former as a synonym for categorical data. However, while nominal data are clearly qualitative, ordinal data are an intermediate type, often treated as quantitative in practice. Moreover, quantitative data can sometimes be treated as categorical. Quantitative data are either *continuous* or *discrete*. Continuous variables can take on any real value in a given interval, whereas discrete variables are restricted to a fixed set of distinct values (e.g., integers). Discrete quantitative variables are often considered as categorical when the range of observed values is small. Count data are an example.

Categorical variables can be divided into *dichotomous* variables that only have two possible outcomes ("yes"/"no"; "male"/"female") and *polytomous* variables with more than two categories ("Christian," "Muslim," "Hindu," "Buddhist," "Jewish," etc.). A dichotomous variable that is coded 0 and 1 is called an *indicator* or *binary* variable. Dichotomous variables are often technically and conceptually easier to handle than polytomous variables.

More of a philosophical question, which also was the root of the dispute between Pearson and Yule, is whether categorical variables should be treated as inherently categorical (Yule's view) or whether they can be conceptualized as coarse-grained manifestations of an underlying continuum (Pearson's position). The distinction, however, has practical consequences for how categorical data are modeled and analyzed. Both views have their justifications, but in some cases, it is more natural to assume fixed categories (for nominal variables such as gender or nationality), whereas in others the assumption of an underlying continuum, or a so-called latent variable, is more applicable (for ordinal data, in particular).

A strong case for the latent variable approach can be made for hybrid variables where a continuous variable is only partially observed. For example, think of a variable classifying households by income; for each household, the range, but not the exact value, of income is known. For such *interval data*, an underlying continuum exists, and thus it is natural to model the data using a latent variable. *Truncated* or *censored variables* (sometimes called *limited variables*) are a related type of hybrid variable, where the continuous data are only observed under certain conditions or within a certain range (e.g., top-coded earnings). A somewhat different type of hybrid data is *time-to-event data* (survival data), where a categorical variable (the type of event) is tied to a quantitative variable—the duration until the event occurs, possibly censored for some observations (e.g., the timing of military interventions in conflicts or the career paths of politicians).

Categorical Data Analysis

Depending on research interest and data type, a variety of categorical data measures and methods are available.

Describing Distributions of Categorical Variables

One may first wish to describe the relative frequencies (proportions) of the different categories of a categorical variable. For example, in the context of election forecasts, one predicts different parties' shares of the vote by polling a random sample of voters. If data are available for the entire population, the proportions can be derived by simple counting. Usually, however, the shares have to be inferred from a random sample. The theoretical distribution used to model statistical properties of dichotomous variables in random samples is the *binomial distribution*, which can be approximated by a normal distribution for large samples. For polytomous variables, the closely related *multinomial distribution* is used. Based on these distributions, exact and approximate (i.e., large sample) methods have been developed to compute *confidence intervals* for proportions and to test whether an observed distribution deviates significantly from a given theoretical distribution of interest (*goodness-of-fit test*). The most common (multinomial) goodness-of-fit test is the *chi-squared test*, which is based on the squared differences between the observed category counts and the counts that are to be expected (on average) according to the theoretical distribution. For ordinal variables, the cumulative distribution function (which is a step function) and tests that exploit this additional information, such as the discrete *Kolmogorov-Smirnov test*, are of interest.

For nominal data, a sensible measure of location is the *mode* (most frequent category). For ordinal variables, the *median* is also useful (the middle of the ordered data), as well as the arithmetic mean if the differences among the categories are assumed to be comparable. A measure of variation for nominal data is, for example, *Simpson's index of diversity*, which reaches its maximum if the distribution is uniform (an equal frequency of each category). For ordinal data, the *range* and measures based on differences between quantiles, such as the *interquartile range* (range of middle half of data), are also meaningful.

Association Among Categorical Variables

For research questions examining relationships between categorical variables, the joint distribution of the variables can be analyzed using a *cross-classification table* (*contingency table*). The cells of such a two-way table report the frequencies or proportions of the different combinations of the variables' categories. It is often useful to condition on row or column totals so that the differences between conditional distributions can easily be identified. For example, think of a table with party preference in rows and ethnicity in columns. Proportions conditioned on column totals (column percent) would then, in each column, reflect the party preference distribution for a specific ethnicity.

To test whether the data provide support for the hypothesis that the conditional distributions across rows or columns are different or, equivalently, whether the two variables are associated with each other, researchers often use the *chi-squared homogeneity or independence test*, which is based on the squared differences between observed cell counts and the cell counts that would be expected if the two variables were unrelated. Since the chi-squared test may be biased in small samples (especially if there are many cells with low expected frequency), exact tests have also been developed (Fisher's exact test for 2 × 2 tables and its generalizations to $n \times m$ tables).

The chi-squared test statistic depends on the number of observations and the size of the table and is therefore not a very useful association measure. However, two popular chi-squared–based measures that can be compared across tables and samples of different size are Pearson's contingency coefficient and Cramér's *V*. A second class of association measures for nominal data is based on the concept of proportional reduction of error (PRE). PRE measures quantify the degree to which knowledge of an observation's value for one variable can be exploited to predict the value of the other (or, more precisely, to which degree the prediction error can be reduced). Different conceptualizations of prediction errors lead to different PRE measures. Examples include Guttman's lambda, Goodman and Kruskal's tau, or the uncertainty coefficient (likelihood ratio index). If variables are strongly associated, then one variable provides a lot of information for the prediction of the other, and the PRE measures are larger. In the case of two dichotomous variables (2 × 2 tables), additional important measures are the odds ratio (the ratio of cross- products of frequencies in the main and secondary diagonal) and the phi coefficient (point correlation coefficient).

The measures discussed so far can be used for any type of categorical variable. For ordinal variables, additional association measures that indicate the direction of a relation have been proposed. Examples are Goodman and Kruskal's gamma, Kendall's tau-*b*, or Somers' *d*, which can be expressed in terms of the numbers of concordant and discordant pairs. An alternative approach is to assign scores that roughly reflect the distances between categories and then compute the linear correlation. Assigning ranks to observations, instead of scores (using midranks for observations in the same category), leads to the rank correlation (Spearman's rho).

Association measures are useful to describe the relation between two variables. A more general framework to uncover association patterns among two or more categorical variables is provided by *loglinear models*. The basic idea of loglinear models is to express the cell counts in contingency tables as a function of the categories defining the cells. The basic independence model contains one parameter for each category of each variable and simply models the marginal distributions of the variables. Association patterns are then captured by additional parameters for combinations of categories across variables (interaction terms). A vast literature on loglinear models was developed from the 1960s through the 1980s, but the methodology seems to be somewhat less in use today.

Other multivariate methods for categorical data include (multiple) *correspondence analysis* and *latent class analysis*, which are similar to principal component analysis and factor analysis but for categorical data. Latent structure models with categorical observed and continuous latent variables are covered by *latent trait analysis* (*item response theory*).

Further special topics in the analysis of the association between categorical variables are *rater agreement* and *segregation*. The degree to which raters agree with respect to a categorical rating or classification can be analyzed using loglinear models, but specialized measures such as Cohen's kappa are also available. Segregation analysis asks how group members are distributed among the values of a categorical variable (occupational sex segregation, racial residential segregation) and has its own specialized measures, such as Duncan and Duncan's index of dissimilarity.

Explaining Categorical Outcomes

Methods for contingency tables are often exploratory and descriptive in the sense that they do not distinguish between dependent and explanatory variables. To analyze research questions that make assertions about causal relations, regression-type statistical methods are required. Explanatory variables in such models can be continuous or categorical. Regression coefficients for binary explanatory variables simply reflect conditional group differences, and categorical predictors with more than two categories (called *factors*) can easily be included in a regression equation by means of indicator variables for the different categories.

First, assume the dependent variable to be *dichotomous* (binary). For example, think of analyzing union membership of employees (1 = *yes*, 0 = *no*) as a function of variables such as sex, age, or political orientation. The expected value of a binary variable is equal to the probability of Outcome 1. As such, standard linear regression methods that model the conditional expectation can be applied. This linear probability model (LPM) has the advantage that regression coefficients can directly be interpreted as partial effects on the probability of Outcome 1. However, the LPM is not always appropriate (e.g., assuming a linear relation between the explanatory variables and the outcome probability is usually unreasonable), and dedicated binary response models such as the logistic regression (logit model) or the probit model have been developed. In the logit model, the log odds, the logarithm of the odds of Outcome 1, are expressed as a linear function of the explanatory variables. In the probit model, the probability of Outcome 1 is linked to the linear predictor through the cumulative normal distribution. Although the interpretation of the coefficients is somewhat more difficult, these models are generally considered more appropriate than the LPM to analyze a binary dependent variable.

Multinomial logit or probit models are generalizations of these models to nominal dependent variables with more than two outcome categories. Multinomial models contain a separate set of coefficients for each outcome, allowing for complex effect patterns of the explanatory variables on the outcome probabilities. Likewise, ordinal logit or probit models are used to analyze ordinal dependent variables, although these models usually contain only one set of coefficients plus a number of cutoff parameters that reflect the latent spacing of the categories. More flexible variants such as the stereotype logistic regression or the generalized ordered logit model are also available. Powerful extensions to multinomial models are choice models that are based on data containing alternative-specific information. The classic example is travel mode choice, with information on characteristics such as costs, comfort, and travel duration for each mode. Such models include the conditional logit and the alternative-specific multinomial probit. A related model is the nested logit, which is used if the alternatives can be grouped according to a nested structure (e.g., restaurants grouped by cuisine).

As mentioned above, one can treat count data as categorical, particularly when counts are low. The most prominent regression procedure for count data is the Poisson regression that models event counts as a Poisson process. However, because the inherent assumption of events' independence is often too restrictive, more flexible models such as the negative binomial regression are generally preferred. Further, an overrepresentation of zeros is often observed in count data, which can be handled by a zero-inflated count model or, depending on the assumed process, a hurdle model that combines a binary response model with a zero-truncated count model. Note that regression models for binary variables and for count variables are both covered in the framework of generalized linear models (GLMs).

Limited dependent variables are analyzed with methods closely related to binary response models. The most well-known is the censored normal regression (tobit model), which is based on the concept of a latent variable with normally distributed errors, as is the probit model. A natural extension are sample selection models (e.g., the Heckman model), which explicitly model the process causing the censoring. Sample selection models have also been developed for the case of a fully categorical dependent variable.

Another closely related field is event history analysis (survival analysis), which is concerned with the explanation of the duration until the occurrence of an event. Time-to-event data are often modeled in terms of the hazard rate: the conditional probability of an event at a certain time point given that the event has not yet occurred. Accelerated failure time models, which are similar to the tobit model, can also be used. Further, discrete-time hazard models are used if the duration is categorical. These models apply a binary response model to the expanded data containing one record per time point for each observation. Another topic in event history analysis is sequence analysis, used to identify and describe typical patterns of event sequences (e.g., career paths of politicians).

Discrimination and classification is a further area in categorical data analysis that has some resemblance to the regression methods discussed. A key difference, however, is that discriminant analysis makes no assertions about direction of causation. It is an exploratory tool to evaluate how well groups can be distinguished with respect to a set of observed characteristics and to derive optimal rules for classifying new observations for which group membership is unknown.

As outlined in this entry, a well-established and diversified toolbox exists for the analysis of categorical data. Moreover, special areas such as categorical panel data and multilevel models, exact inference for binary response and count data models, or nonparametric estimation have made great advances over the past few decades.

Ben Jann
ETH Zurich
Zurich, Switzerland

See also Censored and Truncated Data; Correspondence Analysis; Cross-Tabular Analysis; Event Counts; Event History Analysis; Fuzzy-Set Analysis; Logit and Probit Analyses; Measurement, Levels; Nonlinear Models; Regression

Further Readings

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley.

Huberty, C. J. (1994). *Applied discriminant analysis*. Hoboken, NJ: Wiley.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables.* Thousand Oaks, CA: Sage.

Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using stata* (2nd ed.). College Station, TX: Stata Press.

Powers, D. A., & Xie, Y. (2008). *Statistical methods for categorical data analysis* (2nd ed.). Bingley, UK: Emerald.

Simonoff, J. S. (2003). *Analyzing categorical data.* New York: Springer.

CAUDILLISMO

Caudillismo refers to a system of political-social domination that arose after the independence wars in Spanish America. The caudillo (strongman—from the Latin word *capitellum*, meaning head) was the head of irregular forces who ruled a politically distinctive territory. These forces were governed through an informal system of sustained obedience based on paternalistic relations to the leader, who attained his position as a result of his forceful personality and charisma. Caudillismo as a concept was first used in Spanish America to describe the characteristics of leaders who challenged the authority of the governments arising from the independence process after 1810 and to refer to the political regimes instated by such leaders. In this limited meaning, the notion of caudillismo is a heuristic instrument for analyzing a given historical period that started after the wars of independence and concluded with the emergence of the national states in the second half of the 19th century. John Lynch, one of the historians who made the greatest efforts to define the characteristics and attributes of Latin America's caudillos and their political leadership conditions, stated that caudillismo was the image of society and that caudillos were its creatures. This entry is a review of different interpretations of caudillismo, its origins, and its forms of leadership.

The terms *caudillismo* and *caudillo* continued to be used after the conditions that gave rise to "classical caudillismo" had disappeared. The extension of the notion to encompass any kind of personalized leadership that exercises power in an arbitrary manner within a context of fragility or crisis of the political institutions is not that relevant to critical qualitative research. Caudillismo is used to designate and also stigmatize the governments of "strongmen" with no contextual reference.

Caudillismo After Independence in Spanish America

We will now elaborate on the limited meaning of caudillismo restricted to the framework of the struggles to control power that followed independence in Spanish America. In 1845, the book *Facundo*, written by Domingo Faustino Sarmiento, provided the classical interpretation of caudillismo in Spanish America in the 1800s. From that standpoint, caudillismo is the expression of political barbarism and the antithesis of a government that ensures security, freedom, and ownership rights for a country's inhabitants. Sarmiento's book is a portrait of Facundo Quiroga, the "Tiger of the Plains," a caudillo in the first half of the 19th century. In Quiroga, Sarmiento believes that he sees the incarnation of the antinomy of civilization and barbarism faced by the peoples of the Americas as a result of their revolutionary experience and geography—the desert—which had turned violence into a lifestyle. Quiroga in the Argentine pampas and José Antonio Paéz in the Venezuela prairies represent caudillismo as a system of government and type of political leadership. Physical vigor, spontaneous cruelty, and the rusticity inherent in the rural world they come from can account for the despotism of the regimes they represent. Antonio López de Santa Anna from Mexico appears as the "Attila of civilization" and Juan Manuel de Rosas as the "River Plate Caligula" of Argentina.

Latin American historiography and the contributions of European and U.S. studies revealed practically unknown facets of the caudillismo phenomenon in the postindependence period, which led to a better understanding of an era obscured by myths and legends. The caudillo figures started to appear with more nuances than in the view of their contemporaries, and it was noted that their governments, in many cases, adapted traditional legality within the emergence of a new context.

The origin of caudillismo has been interpreted differently, encompassing factors such as the militarization of politics as a result of the independence wars, the absence of formal rules after the collapse of the colonial order, the ruralization of power, the importance of monarchic tradition, the legacy of authoritarianism and anarchism from the Spaniards, or the characteristics of the village societies. In Venezuela and the River Plate area, where the war against the royal troops was more radical and the whole of the population took arms, caudillismo developed quicker and in a more pronounced manner. Cattle herders and gauchos, layabouts and bandits joined the *Montoneros* (an unruly group of horsemen who fought against Spanish colonization in the 1800s, usually under the federalist caudillos of the provinces outside Buenos Aires), which were truly "informal armies" that got supplies by pillaging and acted under the subordination of the bosses' prestige.

The militarization of politics and society that outlived the battles for independence linked caudillismo to military power and political competition with armed struggles. The caudillo was first a warrior, during wars of liberation, civil wars, and national wars; the caudillo was the strongman who could recruit troops and protect his people, as John Lynch outlined. Studies on Latin American caudillismo in the 1950s considered it a variable of authoritarianism and militarism in response to anarchy. Agustín Gamarra of Peru, known as "Cuzco's black angel," was seen as a prototype of military leaders who, under different guises, dominated politics in postindependence Spanish America. The militias constituted the core of *Gamarrismo* (a militaristic state based on the [invented] tradition of the Incas). As pointed out by Charles Walker, the militias controlled local society by monitoring any threat of opposition, and they also served as veritable military academies and vehicles of social mobility. High-ranking positions in the militias represented important avenues for political and economic gain. In Mexico and Peru, professional military men played an important role in the political process as pressure groups. In other countries, the military organization of the end of the colonial period was swept away by the wars of independence, although with different outcomes. Nonetheless, some military heads were prevailing figures—for instance, Francisco de Paula Santander in Nueva Granada (present-day Colombia), Juan José Flores in Ecuador, and Paéz in Venezuela and Andrés de Santa Cruz in Bolivia. As Walker notes, the analysis of caudillismo needs to shift its attention from the battlefield and to reconsider the state, as middling officials such as subprefects and militia leaders proved more important for Gamarra than military officers and guerrillas.

Caution is needed with respect to generalizations about Spanish America's politics during the half-century following independence because of the distinctive institutional legacies of colonialism, as well as differences in geographical and ethnic composition between countries where the indigenous population prevails and is only partly assimilated into Spanish culture and countries with a Creole population mainly integrated into Spanish culture. The church and the army were two big corporations during colonial times, with an institutional and economic preponderance in Mexico not found in other regions. This was an important source of divergence between Catholics and anticlericals in Mexico. Conflicts between rural and urban areas, between federalism and centralism, and between rival trends in trade policies—"protectionism" and "liberalism"—all hindered the consolidation of a stable national policy. The duration and pace of the transition to postcolonial order varied between countries, as did the role of the regional or national caudillos. In all cases, however, the constant feature is loyalty to caudillos, most of whom had no defined ideologies and fluctuated between contradictory positions, whether unitary or federal, conservative or liberal, or favoring protectionism or free trade.

Tulio Halperín Donghi explored the relationship between militarization and democratization that is at the heart of the promotion of the caudillo power. Social promotion opportunities arose after the rupture of colonial order. Agustín de Iturbide, the "constitutional emperor of Mexico," belonged to a poor family; Gamarra and Ramón Castilla, both from Peru, were Creoles. They all reached positions that would have before been inaccessible to them, but this relative liberal opening was an instrument to impede the excessive dissemination of popular participation within a context in which power legitimacy was always questioned.

In the middle of the 1970s, Eric Wolf and Edgar Hansen defined caudillismo as a subcategory of "clientelism." These authors believe that caudillismo refers to the quest for power and wealth through violent conquest that establishes a link between protector and protected in an unstable society lacking the institutional channels through which people can compete for these social and financial advantages. John Lynch also states that the patron–client relationship was the essential link that was finally built into the state and became the model of caudillismo. Alan Knight's more recent view underlines the relevance of traditional authority in village societies, combined with elements of charisma, to understand the command–obedience relationship inherent in caudillismo.

The notion of caudillismo encompasses different leadership subcategories, for instance, the "sword caudillos," the warlords who dominated during a period of great instability when popular support and force were decisive. Pacification and progressive institutionalization of power relations were carried out by new leader profiles, the "tame caudillos," who were deemed to be the arbitrators capable of achieving national unification and building more stable institutions. Rosas's government (1829–1852) is an example of a new modality. Sarmiento believed that was the end of caudillismo since it combined elements of barbarism with elements of civilization. Cruelty became more sophisticated, the press and the legislature replaced barbarian passion, and strong personalism allowed for the exertion of influence on the political orientation of other River Plate provinces, although Rosas's powers were not included in a constitutional system. In Mexico, Benito Juárez, a lawyer born into a poor family, used the 1857 Constitution to take on emergency powers and govern in an authoritarian manner, constitutional but cruel. Diego Portales's constitutional authoritarianism sought support, as he himself defined, in "the weight of night," in the acquiescence of the masses, passive and ignorant. A prevailing political figure in Chile in the 1830s, Portales was a trader and omnipotent minister but never took office as president. From Sarmiento's viewpoint, Portales was an example of an illustrated caudillo. The use of the term *caudillo* was extended to encompass authoritarian leaders such as Juárez or Portales, who governed within the framework of constitutions. According to Lynch, the "impenetrable dictatorship" of Dr. José Francia in Paraguay combined despotism and isolation for 3 decades in which he ensured the basic needs of a Creole society.

The notion of caudillismo makes possible a better understanding of the political processes that took place during that period since interpretation of the past cannot ignore the categories used by their contemporaries.

Liliana De Riz
Universidad de Buenos Aires
Buenos Aires, Argentina

See also Anarchy; Charisma; Clientelism; Regionalism; State Formation

Further Readings

Knight, A. (1980). Peasant and caudillo in revolutionary Mexico 1910–1917. In D. A. Brading (Ed.), *Caudillo and peasant in the Mexican revolution* (pp. 17–58). Cambridge, UK: Cambridge University Press.

Lynch, J. (1992). *Caudillos in Spanish America, 1800–1850*. Oxford, UK: Clarendon Press.

Tulio, H. D. (1993). *The contemporary history of Latin America*. Durham, NC: Duke University Press.

Walker, C. F. (1999). *Smoldering ashes: Cuzco and the creation of republican Peru 1780–1840*. Durham, NC: Duke University Press.

Wolf, E. R., & Hansen, E. C. (1967). Caudillo politics: A structural analysis. *Comparative Studies in Society and History, 9,* 168–179.

CAUSAL HETEROGENEITY

See Causality

CAUSALITY

Causality refers to the relationship between events where one set of events (the effects) is a direct consequence of another set of events (the causes). Causal inference is the process by which one can use data to make claims about causal relationships. Since inferring causal relationships is one of the central tasks of science, it is a topic that has been heavily debated in philosophy, statistics, and the scientific disciplines. This entry reviews the models of causation and tools for causal inference most prominent in the social sciences, including regularity approaches, associated with David Hume, and counterfactual models, associated with Jerzy Splawa-Neyman, Donald Rubin, and David Lewis, among many others. One of the most notable developments in the study of causation is the increasing unification of disparate methods around a common conceptual and mathematical language that treats causality in counterfactual terms—that is, the Neyman-Rubin model. This entry discusses how counterfactual models highlight the deep challenges involved in making the move from correlation to causation, particularly in the social sciences, where controlled experiments are relatively rare.

Regularity Models of Causation

Until the advent of counterfactual models, causation was primarily defined in terms of observable phenomena. It was the philosopher Hume in the

18th century who began the modern tradition of regularity models of causation by defining causation in terms of repeated "conjunctions" of events. In *An Enquiry Into Human Understanding* (1751), Hume argued that the labeling of two particular events as being causally related rested on an untestable metaphysical assumption. Consequently, Hume (1739) argued that causality could be adequately defined only in terms of empirical regularities involving classes of events. He asked, "How could we know that a flame caused heat?"—only by calling "to mind their constant conjunction in all past instances. Without further ceremony, we call the one cause and the other effect and infer the existence of one from that of the other" (*Treatise of Human Nature*, Book 1, Pt. 3, sec. 6). In the *Enquiry*, Hume argued that three empirical phenomena were necessary for inferring causality: (1) contiguity (the cause and effect must be contiguous in time and space), (2) succession (the cause must be prior to the effect), and (3) constant conjunction (there must be a constant union between the cause and effect). Under this framework, causation was defined purely in terms of empirical criteria, rather than unobservable assumptions. In other words, Hume's definition of causation and his mode of inference were one and the same.

John Stuart Mill, who shared the regularity view of causation with David Hume, elaborated basic tools for causal inference that were highly influential in the social sciences. For Mill, the goal of science was the discovery of regular empirical laws. To that end, Mill proposed in his 1843 *A System of Logic*, a series of rules or "canons" for inductive inference. These rules entailed a series of research designs that examined whether there existed covariation between a hypothesized cause and its effect, time precedence of the cause, and no plausible alternative explanation of the effect under study. Mill argued that these research designs were effective only when combined with a manipulation in an experiment. Recognizing that manipulation was unrealistic in many areas of the social sciences, Mill expressed skepticism about the possibility of causal inference for questions not amenable to experiments.

The most widely used of Mill's canons, the "Direct Method of Difference," entailed the comparison of two units identical in all respects except for some manipulable treatment. The method of difference involves creating a counterfactual control unit for a treated unit under the assumption that the units are exactly alike prior to treatment, an early example of counterfactual reasoning applied to causal inference. Mill (1884) stated the method as follows:

> If an instance in which the phenomenon . . . occurs and an instance in which it does not . . . have every circumstance save one in common . . . [then] the circumstance [in] which the two instances differ is the . . . cause or a necessary part of the cause. (Book 3, chap. 8)

The weakness of this research design is that in practice, particularly in the social sciences, it is very difficult to eliminate all heterogeneity in the units under study. Even in the most controlled environments, two units will rarely be the same on all background conditions. Consequently, inferences made under this method require strong assumptions.

Mill's and related methods have been criticized on a variety of grounds. His canons and related designs assume that the relationship between cause and effect is *unique* and deterministic. These conditions allow neither for more than one cause of an effect nor for interaction among causes. The assumption that causal relationships are deterministic or perfectly regular precludes the possibility of measurement error. If outcomes are measured with error, as they often are in the social sciences, then methods predicated on detecting constant conjunctions will fail. Further, the causal relationships typically studied in the social and biological sciences are rarely, if ever, unique. Causes in these fields are more likely to have highly contingent effects, making regular causal relationships very rare.

Counterfactual Models of Causation

Regularity models of causation have largely been abandoned in favor of counterfactual models. Rather than defining causality purely in reference to observable events, counterfactual models define causation in terms of a comparison of observable and unobservable events. Linguistically, counterfactual statements are most naturally expressed using subjunctive conditional statements such as "If India had not been democratic, periodic famines would have continued." Thus, the counterfactual approach to causality begins with the idea that some of the information required for inferring causal relationships is and always will be unobserved, therefore some assumptions must be made. In stark contrast to the regularity approach of Hume, the fact of counterfactual causation is fundamentally separate from the tools used to infer it. As a result, philosophers like Lewis (1973) could write about the meaning of causality with little discussion of how it might be inferred. It was statisticians, beginning with Splawa-Neyman in 1923 and continued most prominently by Rubin, who began to clarify the conditions under which causal inferences were possible if causation was fundamentally a "missing-data problem."

Counterfactual Models in Philosophy

In philosophy, counterfactual models of causation were largely absent until the 1970s due to Willard van Orman Quine's dismissal of the approach in his *Methods of Logic* (1950), where he pointed out that counterfactual statements could be nonsensical. He illustrated this point by his famous comparison of the conditional statements "If Bizet and Verdi had been compatriots, Bizet would have been Italian" and "If Bizet and Verdi had been compatriots, Verdi would have been French." For Quine, the incoherence of the two statements implied that subjective conditionals lacked clear and objective truth conditions. Quine's suspicion of conditional statements was also rooted in his skepticism of evaluating the plausibility of counterfactual "feigned worlds," as he explained in *Word and Object* (1960):

> The subjunctive conditional depends, like indirect quotation and more so, on a dramatic projection: we feign belief in the antecedent and see how convincing we then find the consequent. What traits of the real world to suppose preserved in the feigned world of the contrary-to-fact antecedent can only be guessed from a sympathetic sense of the fabulist's likely purpose in spinning his fable. (p. 222)

Perhaps because of this view of counterfactuals, Quine had a dim view of the concept of causality. He argued that as science advanced, vague notions of causal relationships would disappear and be replaced by Humean "concomitances"—that is, regularities.

In philosophy, Lewis popularized the counterfactual approach to causality 50 years after it first appeared in statistics with Splawa-Neyman's 1923 paper on agricultural experiments. For Lewis, Quine's examples revealed problems only with *vague* counterfactuals, not counterfactuals in general. A cause, according to Lewis in his 1973 article "Causation," was "something that makes a difference, and the difference it makes must be a difference from what would have happened without it" (p. 557). More specifically, he defined causality in terms of "possible" (counterfactual) worlds. He postulated that one can order possible worlds with respect to their closeness with the actual world. Counterfactual statements can be defined as follows:

> If *A* were the case, *C* would be the case is true in the actual world if and only if (i) there are no possible *A*-worlds; or (ii) some *A*-world where *C* holds is closer to the actual world than is any *A*-world where *C* does not hold. (p. 560)

More intuitively, causal inferences arise by comparing the actual world with the closest possible world. If *C* occurs in both the actual and the closest possible world without *A*, then, according to Lewis, *A* is not the cause of *C*. If, on the other hand, *C* does not occur in the closest possible world without *A*, then *A* is a cause of *C*. Lewis's theory was concerned with ontology, not epistemology. As a result, one might argue that his work has limited use to empirical research since he provided little practical guidance on how one could conjure the closest possible worlds to use as comparison cases. Without additional assumptions, Lewis's model suggests that causal inference is a fruitless endeavor given our inability to observe nonexistent counterfactual worlds.

Statistical Models of Causation

Fortunately, statisticians beginning with Splawa-Neyman in 1923 elaborated a model of causation that allowed one to treat causation in counterfactual terms and provided guidance on how empirical researchers could create observable counterfactuals.

Say we are interested in inferring the effect of some cause T on a parameter $\bar{Y}$ of the distribution of outcome Y in population A relative to treatment C (control). Population A is composed of a finite number of units and $\bar{Y}_{A,T}$ is simply a summary of the distribution of that population when exposed to T, such as the mean. If treatment C (control) were to be applied to population A, then we would observe $\bar{Y}_{A,C}$. To use Lewis's terminology, in the actual world, we observe $\bar{Y}_{A,T}$, and in the counterfactual world, we would observe $\bar{Y}_{A,C}$. The causal effect of T relative to C for population A is a measure of the difference between $\bar{Y}_{A,T}$ and $\bar{Y}_{A,C}$, such as $\bar{Y}_{A,T} - \bar{Y}_{A,C}$. Of course, we can observe only the parameter that summarizes the actual world and not the counterfactual world.

The key insight of statistical models of causation is that under special circumstances we can use another population, B, that was exposed to control to act as the closest possible world of A. If we believe that $\bar{Y}_{A,C} = \bar{Y}_{B,C}$, then we no longer need to rely on an unobserved counterfactual world to make causal inferences; we can simply look at the difference between the observed $\bar{Y}_{A,T}$ and $\bar{Y}_{B,C}$. In most cases $\bar{Y}_{A,C} \neq \bar{Y}_{B,C}$, however, so any inferences made by comparing the two populations will be *confounded*. What are the special circumstances that allow us to construct a suitable counterfactual population and make unconfounded inferences? As discussed below, the most reliable method is through randomization of treatment assignment, but counterfactual inferences with observational data are possible—albeit more hazardous—as well. In either case, causes are defined in reference to some real or imagined *intervention*, which makes the counterfactuals well-defined.

The Neyman-Rubin Model

The counterfactual model of causation in statistics originated with Splawa-Neyman's 1923 model, which is nonparametric for a finite number of treatments, where each unit has a potential outcome for each possible treatment condition. In the simplest case with two treatment conditions, each unit has two potential outcomes, one if the unit is treated and the other if untreated. In this case, a causal effect is defined as the difference between the two potential outcomes, but only one of the two potential outcomes is observed. In the 1970s, Rubin developed the model into a general framework for causal inference with implications for observational research. Paul Holland in 1986 wrote an influential review article that highlighted some of the philosophical implications of the framework. Consequently, instead of the "Neyman-Rubin model," the model is often simply called the Rubin causal model or sometimes the Neyman-Rubin-Holland model or the Neyman-Holland-Rubin model.

The Neyman-Rubin model is more than just the math of the original Neyman model. Unlike Splawa-Neyman's original formulation, it does not rely on an urn model motivation for the observed potential outcomes but rather on the random assignment of treatment. For observational studies, one relies on the assumption that the assignment of treatment can be treated as if it were random. In either case, the mechanism by which treatment is assigned is of central importance. The realization that the primacy of the assignment mechanism holds true for observational data no less than for experimental data is due to Rubin. This insight has been turned into a motto: "No causation without manipulation."

Let Y_{iT} denote the potential outcome for unit i if the unit receives treatment, and let Y_{iC} denote the potential outcome for unit i in the control regime. The treatment effect for observation i is defined by $\tau_i = Y_{iT} - Y_{iC}$. Causal inference is a missing-data problem because Y_{iT} and Y_{iC} are never both observed. This remains true regardless of the methodology used to make inferential progress—regardless of whether we use quantitative or qualitative methods of inference. The fact that we cannot observe both potential outcomes at the same time is commonly referred to as the "fundamental problem of causal inference." Let T_i be a treatment indicator: 1 when i is in the treatment regime and 0 otherwise. The observed outcome for observation i is then

$$Y_i = T_i Y_{iT} + (1 - T_i) Y_{iC}.$$

The average causal effect τ is the difference between the expected values $E(Y_T)$ and $E(Y_C)$. We observe only the conditional expectations $E(Y_T|T = 1)$ and $E(Y_C|T = 0)$, not the unconditional expectations required for obtaining τ. Until we assume that $E(Y_T|T = 1) = E(Y_T)$ and

$E(Y_C|T = 0) = E(Y_C)$, we cannot calculate the average treatment effect. Note that the estimand of interest, such as the average treatment effect, is conceptually distinct from the estimators used to infer it from data, such as difference in means, linear regression, or other techniques.

Experiments

To estimate the average treatment effect, we require the assumption of *independence*. The singular virtue of experiments is that physical randomization of an intervention ensures independence between treatment status and potential outcomes. Ronald Fisher, in the 1920s and 1930s, first emphasized the importance of random assignment for eliminating bias, calling randomization of treatment the "reasoned basis for inference." From a Lewisian perspective, the control group in an experiment functions as an observable "possible world." With the independence assumption, the average treatment effect can be estimated from observables using the following expression:

$$\begin{aligned}\tau &= E(Y_{iT}|T = 1) - E(Y_{iC}|T = 0)\\ &= E(Y_{iT}) - E(Y_{iC}).\end{aligned}$$

Under randomization, the assumption that T_i is independent of Y_{iT} and Y_{iC} is plausible, making the treatment and control groups exchangeable in expectation.

One of the assumptions that randomization by itself does not justify is that the response of one unit should be unaffected by the particular assignment of treatments to the other units. This "no interference between units" is often called the Stable Unit Treatment Value Assumption (SUTVA). SUTVA implies that the potential outcomes for a given unit do not vary with the treatments assigned to any other unit and that there are not different versions of treatment.

Observational Data

In observational data, stronger assumptions are usually required to estimate causal effects. In observational studies, the causal quantity of interest is often the "average treatment effect on the treated" or ATT, which is the average effect conditional on being in the treatment regime. The parameter of interest is

$$\tau|(T = 1) = E(Y_{iT}|T = 1) - E(Y_{iC}|T = 1).$$

Since the counterfactual control units, $E(Y_{iC}|T = 1)$, are not observed, a control group must be constructed. The two assumptions required to construct a valid control group are conditional independence of the potential outcomes and treatment assignment and overlap, or

1. $Y_{iT}, Y_{iC} \perp T|X$,
2. $0 < \Pr(T = 1|X) < 1$.

When these two conditions hold, we can say that treatment assignment is *strongly ignorable*. Once a control group is constructed that enables us to satisfy these two conditions, the ATT can be estimated as

$$\begin{aligned}t|(T = 1) = E[E(Y_{iT}|T = 1)\\ - E(Y_{iC}|T = 0)]|T = 1.\end{aligned}$$

It is important to note that the outer expectation is taken over the distribution of $X|(T = 1)$, which is the distribution of covariates among the treated units.

Note that the ATT estimator is changing when individual observations are weighted, and the observations that are outside of common support receive zero weights. That is, if some covariate values are observed only for control observations, those observations will be irrelevant for estimating ATT and are effectively dropped. Therefore, the overlap assumption for ATT requires only that the support of X for treatment observations be a subset of the support of X for control observations. More generally, one would also want to drop treatment observations if they have covariate values that do not overlap with control observations. In such cases, it is unclear exactly what estimand one is estimating because it is no longer ATT as some treatment observations have been dropped along with some control observations.

The key assumption being made here is strong ignorability. Even thinking about this assumption presupposes some rigor in the research design. For example, is it clear what is pretreatment and what is posttreatment? If not, one is unable even to form the relevant questions, the most useful of which may be the one suggested by Harold Dorn in 1953, who proposed that the designer of every observational study should ask, "How would the study be conducted if it were possible to do it by controlled experimentation?" This clear question also appears in William Cochran's 1965 Royal Statistical Society discussion paper on the planning of observational studies of human populations. Dorn's question has become one that researchers in the tradition of the Neyman-Rubin model ask themselves and their students. The question forces the researcher to focus on a clear manipulation and then on the selection problem at hand. Only then can one even begin to think clearly about how plausible the strong ignorability assumption may or may not be.

Structural Equation Modeling

Another prominent approach to causal inference using counterfactuals is structural equation modeling, a method most associated with the work of Judea Pearl. Structural equation modeling is an old enterprise that has a rich history, including foundational work on causality in systems of structural equations by the geneticist Sewall Wright (1921), the economist Trygve Haavelmo (1943), and the political scientist Herbert Simon (1953). Modern advocates of structural equation modeling argue that the probability calculus approach to causal modeling used by researchers in the Neyman-Rubin tradition is too narrow in that it does not explicitly take into account knowledge about the mechanisms linking background, independent variables, and dependent variables. Rather than modeling causality in relation to experiments, structural equation modelers prefer to write out a more elaborate causal model of the relationships under investigation through a system of structural functions. A system of such functions is said to be structural if the functions are assumed to be invariant to possible changes in the form of the other functions. Under this framework, the effects of treatments are understood as interventions in a prespecified structural model.

For structural equation modelers, hypothetical interventions should be explicitly and formally related to the causal mechanisms under study. In Pearl's version of structural equation modeling, for example, the mathematical operator "*do(X)*"is used to represent physical interventions in a set of equations that deletes certain functions from the model, replaces them by a constant, and preserves the rest of the model. The counterfactual conditional "If X had been x" is interpreted as an instruction to modify the original model so that some causal variable X is set to x by some intervention, experimental or otherwise. This operator is accompanied by a set of rules called "do calculus," which helps a researcher judge whether or not sufficient information exists to identify the effect of the intervention of interest. Rather than identifying one all-encompassing assumption—strong ignorability—as in the Neyman-Rubin approach, Pearl proposes that researchers adopt a series of local assumptions about how an intervention interacts with a prespecified structural model to identify causal quantities. Despite the rather substantial conceptual differences between these two approaches, however, they are mutually compatible. This compatibility arises from their shared reliance on counterfactual understandings of causality.

Causal Mechanisms

The Neyman-Rubin counterfactual approach is primarily concerned with defining *what* the effect of a cause is, not explaining *how* causes affect outcomes. The apparatus of most statistical models of causation has no formal role for social theory, explanation, or causal mechanisms. Given social scientists' interest in these issues, a common critique of the Neyman-Rubin model and its cousins are that they are too narrow for the social sciences. Advocates of the statistical approach have countered that counterfactual models of causation can be augmented to take into account the causal mechanisms.

While experiments have the virtue of credibly identifying the causal effect of an intervention, they are sometimes criticized as "black boxes." To understand the pathways by which interventions affect the outcome, social scientists have relied on a method known as "mediation analysis," which models the relationship between a treatment, a potentially posttreatment variable, and the outcome ultimately of interest. An important distinction in this literature is whether or not the "mediator" is treated as posttreatment or not. If the mediator is not affected by treatment, the effect of interest is how the manipulable mediator affects or moderates the outcome when the main treatment variable is fixed, known as the "controlled direct effect."

This controlled direct effect is not always the effect of interest, however, since mediation analysis is often intended to shed light on the role of mechanisms, which in this framework can be defined as a process that can transmit, at least partially, the effect of a treatment on an outcome. An important distinction between a manipulation (a "treatment") and a mechanism is that the former involves an external intervention, while the latter does not. The goal of this type of mediation analysis is to estimate what fraction of a causal effect is "indirect"—that is, due to the treatment changing the level of the mediator and consequently the outcome—and what fraction is "direct"—that is, due to the treatment affecting the outcome through other pathways. Expressed in counterfactual language, an "uncontrolled" indirect effect is a comparison between the outcome when the mediator is set at the value realized in the treatment condition and the outcome when the mediator is set to the value that would be observed under the control condition while holding treatment status constant.

Uncontrolled mediation effects are often of great interest, but unfortunately, even with a randomized intervention, their identification rests on strong assumptions. In mediation analyses, the level of the mediator is generally assumed to be independent of the counterfactual outcomes conditional on treatment assignment—that is, the mediator is assigned "as if" random. Given that an uncontrolled mediator variable, by definition, is not randomly assigned, this assumption is strong indeed. While the identification assumptions may be warranted in special circumstances, the main lesson of the statistical literature is that the quantitative study of causal mechanisms is an enterprise fraught with difficulties, even in the context of randomized experiments.

Qualitative Evidence and Theory Falsification

While the quantitative study of causality is well developed and increasingly unified under counterfactual models, many social scientists supplement statistical methods with qualitative reasoning to aid causal inference. Sometimes called "causal process observations," qualitative evidence can be an important source of leverage for both the design of causal analyses and the interpretation of their findings. In the social sciences, for example, most evidence for mechanisms is qualitative, not quantitative. Qualitative researchers argue that by direct observation of causal processes, a researcher can discern potentially important mechanisms that may have escaped notice. Insight derived from observations that are poorly suited for rectangular data sets may then lead to more formal investigations using experimental and observational quantitative methods. The health sciences, for example, are replete with examples of qualitative observation paving the way for groundbreaking experiments.

Given that most questions in the social sciences are studied using observational research designs, another role for qualitative insight is the justification of the conditional independence assumption. Although many and perhaps most observational studies pay inadequate attention to justifying the adequacy of their designs, careful observational research must identify important confounders and uncover fortuitous "natural" experiments for making well-grounded inferences. Qualitative evidence can be used to identify appropriate confounders to adjust for as well as to justify any claim that treatment was allocated "as if" random.

For many questions in the social sciences, however, a research design guaranteeing the validity of causal inferences is difficult to obtain. When this is the case, researchers can attempt to defend hypothesized causal relationships by seeking data that subject their theory to repeated falsification. Karl Popper famously argued that the degree to which we have confidence in a hypothesis is not necessarily a function of the number of tests it has withstood but rather the severity of the tests to which the hypothesis has been subjected. A test of a hypothesis with a design susceptible to hidden bias is not particularly severe or determinative. If the implication is tested in many contexts, however, with different designs that have distinct sources of bias, and the hypothesis is still not rejected, then one may have more confidence that the causal relationship is genuine. Note that repeatedly testing a hypothesis with research designs suffering from similar types of bias does not constitute a severe test, since each repetition will merely replicate the biases of the original design. In cases where randomized experiments are infeasible or credible natural experiments are unavailable, the inferential difficulties facing researchers are large. In such circumstances, only creative and severe falsification tests can make the move from correlation to causation convincing.

F. Daniel Hidalgo and Jasjeet S. Sekhon
University of California, Berkeley
Berkeley, California, United States

See also Experiments, Field; Experiments, Natural; Matching; Quantitative Methods, Basic Assumptions; Statistics: Overview

Further Readings

Brady, H. E. (2008). Causation and explanation in social science. In J. Box-Steffensmeier, H. Brady, & D. Collier (Eds.), *The Oxford handbook of political methodology* (pp. 217–270). New York: Oxford University Press.
Fisher, R. A. (1935). *The design of experiments.* Edinburgh, UK: Oliver & Boyd.
Greenland, S., Robins, J. M., & Pearl, J. (1999). Confounding and collapsibility in causal inference. *Statistical Science*, *14*(1), 29–46.
Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica, 11*(1), 1–12.
Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945–960.
Hume, D. (1902). *Enquiries concerning the human understanding and concerning the principles of morals* (2nd ed.; L. A. Selby-Bigge, Ed.). Oxford, UK: Clarendon.
Hume, D. (2000). *A treatise of human nature* (Oxford Philosophical Texts). New York: Oxford University Press. (Origina work published 1739)
Lewis, D. K. (1973). Causation. *Journal of Philosophy, 70*(17), 556–567.
Lewis, D. K. (2001). *Counterfactuals.* Malden, MA: Wiley-Blackwell.
Mill, J. S. (1884). *A system of logic, ratiocinative and inductive.* London: Longmans.
Pearl, J. (2000). *Causality.* New York: Cambridge University Press.
Popper, K. (1959). *The logic of scientific discovery.* London: Routledge.
Quine, W. V. O. (1974). *Methods of logic.* London: Taylor & Francis.
Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology, 3*(2), 143–155.
Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*(5), 688–701.
Rubin, D. B. (2006). *Matched sampling for causal effects.* New York: Cambridge University Press.
Splawa-Neyman, J. (1990). On the application of probability theory to agricultural experiments: Essay on principles. Section 9. (D. M. Dabrowska & T. P. Speed, Trans.). *Statistical Science, 5*(4), 465–472. (Original work published 1923)

Censored and Truncated Data

Both censored and truncated data involve a lack of information about a random variable and occur in the context of quantitative analysis of data, when one is using that variable either to estimate a population mean (or other population parameters) or as the dependent variable in a regression analysis. The key distinction between them is whether one has information about missing values. With censored data, one observes some information about the missing data, either in the form of a range of values that they might fall into or in the form of the knowledge that they are missing. With truncated data, one has no information about the existence or value of missing observations. The difference in the structure of information for two types of data determines how one approaches censored or truncated data, whether in the context of a single random variable or as a dependent variable in a regression analysis. This entry discusses the consequences of this kind of missing data for regression analysis and sample selection.

Concerns about censoring and truncation abound in empirical analysis. They can occur either through the structure of data-gathering efforts or through legal requirements. Historically, researchers relied on assumptions about the distribution of the variable to adjust the estimates to account for truncation or censoring. These concerns are just as important when censored or truncated variables are the dependent variables in regression analyses. Scholars have long worried about the consequences of processes such as self-selection for the validity of their regression results, since they often result in biased and inconsistent coefficient estimates. While estimators to correct for censoring and truncation have been around for more than 3 decades, researchers have worried about their sensitivity to distributional assumptions and model specification. While recent work attempts to relax some of these critical assumptions and diagnose sensitivity issues, researchers have also extended previous work by designing estimators for a greater variety of data.

Types of Missing Data

Censored and truncated data are both forms of missing data, which have been categorized in three ways: (1) missing at random (MAR), (2) missing completely at random (MCAR), and (3) nonignorable (NI). A variable that is MCAR has missing values that are determined randomly, so that they occur with equal probability and do not depend on any information in the data set. A variable that is MAR exhibits a pattern of missingness in which the probability of a missing value depends on other observed variables for that same observation. A variable exhibits NI missingness when unobserved information in its value helps explain its missingness. In this case, the pattern of missingness depends on information beyond that which is contained in the observed variables.

Censored and truncated data can emerge through any of these three forms of missing data. Making valid inferences requires making valid assumptions about the structure of the missingness. Except for the case of MCAR, one will generally reach inaccurate conclusions, whether regarding population characteristics or regression parameters, unless one properly models the pattern of missingness.

Censored and Truncated Random Variables

A random variable is censored if one does not observe its true value but rather observes a bound for the range of values into which it falls. If X is a random variable, then X is censored if when $X > b$ the researcher observes only b. Similarly, X may be censored from below by a value a. The density function of the observed values of X therefore has point masses at the values a and b. This form of censoring occurs commonly in political science in the study of the duration of political events. For example, if a researcher observes a set of states that might adopt a particular policy, then states that have not adopted the said policy by the end of the study period are said to be right censored. For those states, the researcher does not know exactly how many years will elapse before they adopt the policy, only that it exceeds the number of years in the study period so far.

A random variable is truncated if the researcher has no information about missing values, including their existence. For example, observations with $X < a$ might not be observed at all. Observations of truncated data do not exhibit point masses at the truncation points. Rather, one accounts for the missing values by focusing on the distribution conditional on observation. This involves normalizing the distribution of the underlying variable, which includes observed and unobserved values, by the probability of the observed region.

An example occurs in the study of campaign finance in the United States: Political contributions to candidates for national office are not reported if the amount does not exceed \$200. Data from the government therefore list no contributions less than \$200, making it impossible to know the number of such contributions or their amount if any such contributions were made.

Censoring and Truncation in Regression Analysis

Regression analysis of censored and truncated variables often requires adaptations of the standard regression model. If the dependent variable is censored, then one must treat observations that take on the censoring value differently (note that the value at which censoring occurs and the value that is observed for censored observations need not be identical). In the case of state policy adoption, discussed above, the data are censored from above, in the sense that the dependent variable for states that have not adopted the said policy by the end of the study period takes on a value equal to the length of the study. In the context of duration analysis, this is referred to as right censoring. If one treats these observations the same way as those observations that have actually been adopted in the past year, bias will likely ensue. Rather, one should model the information accurately: One knows only that the observation is greater than the censoring value.

Another common form of censoring occurs when smaller values of the dependent variable are not observed. Often this corresponds to situations in which the dependent variable takes on nonnegative values, such as one country's foreign aid to other countries. Many countries will receive no foreign aid dollars, while others will receive positive amounts. The clump of countries receiving zero dollars can be thought of as cases in which the host country might have preferred to give negative aid, but since this is impossible, the observed value is censored at zero. One would generally analyze data with this structure with a Tobit model, which accounts for the censoring process.

When the dependent variable suffers from truncation, one does not have information about data for which the dependent variable is outside the bound(s) of truncation. To account for this structure, one must normalize the distribution of the dependent variable to account for the lack of observations in certain regions (so that the distribution over observed values integrates to one). In the case of campaign contributions discussed above, one would observe information only about contributors that donate at least \$200, since that is the legal cutoff for reporting requirements. Thus, one would likely have a data set with complete information for all individuals who contribute at least \$200 but no information and no observations for individuals who contribute less than that amount. Unlike the case of censoring, however, there will not be a large number of individuals clustered at \$200.

Censoring and Truncation Through Sample Selection

A related form of censoring and truncation occurs when a second process determines whether the dependent variable is missing for a given observation. This is often referred to as nonrandom sample selection and can result in selection bias if one ignores the selection process. One can think of the data-generating process as occurring in two stages: The selection equation determines whether one observes the dependent variable of interest for an observation, and the equation of interest determines the value of this variable given the observation. When the unobserved component of the selection equation is correlated with the unobserved component in the outcome equation, nonrandom sample selection is a form of nonignorable missingness. This will generally result in biased coefficients if one ignores the selection process.

Nonrandom sample selection can take two forms. If one observes cases for which the dependent variable is missing, the data suffer from stochastic censoring. One knows that the observation could have experienced the outcome of interest, but because of the selection process, it happened not to experience it. Alternatively, one might have information only about observations with observed values of the outcome of interest. For example, if a survey respondent answers all questions save the dependent variable, then the data exhibit stochastic censoring; whereas if a potential respondent chooses not to answer the entire survey, then the data exhibit stochastic truncation.

Both can be dealt with in the regression context by simultaneously modeling the selection process and the outcome of interest. The most well-known estimator of this type was developed by James Heckman and is often referred to as the Heckman model. While the Heckman model is designed for a continuous dependent variable and data that suffer from stochastic censoring, similar estimators exist for discrete outcomes. Analogous estimators also exist for continuous and discrete outcome variables when the data exhibit stochastic truncation.

Estimation requires proper specification of the selection equation. Furthermore, the estimates tend to be sensitive to the distributional assumption for the error terms (generally bivariate normal). Dealing with stochastic truncation generally poses a greater estimation challenge because one has no information about incomplete observations, making the distributional assumption even more important, since it is the mechanism for identifying the selection process.

Frederick J. Boehmke
University of Iowa
Iowa City, Iowa, United States

See also Data, Missing; Regression; Sampling, Random and Nonrandom; Selection Bias

Further Readings

Achen, C. H. (1986). *The statistical analysis of quasi-experiments*. Berkeley: University of California Press.
Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153–161.
Maddala, G. G. (1983). *Limited dependent and qualitative variables in econometrics*. Cambridge, UK: Cambridge University Press.
Winship, C., & Mare, R. D. (1992). Models for sample selection bias. *Annual Review of Sociology, 18,* 327–350.

CENSORSHIP

Censorship is restriction of the circulation of information, ideas, and images or of political, cultural, religious, and artistic opinions; it occurs mainly when the rulers consider that these could weaken or damage their hold on power. Characteristic of nondemocratic systems, and also present in some democracies, censorship is usually exercised by an authority responsible for monitoring the dissemination of ideas, including those circulated by governmental agencies as well as those from private individuals and organizations. The term *censorship* has its origin in the Roman Empire, where the "censor" was the person who controlled the dissemination of political ideas.

Since premodern times, censorship has existed under different political regimes. It began to be questioned in the 17th century, when critics of the period's absolutist states started to call for the limitation of state power and discretion and for the institutionalization of civil and political rights and, particularly, freedom of speech.

Censorship has also existed in social institutions such as the Catholic Church as well as in other religions such as Islam. In this case, the upper echelons of the religious hierarchy exercise censorship to ensure that the content of writings by priests and bishops is in line with the religion's main tenets.

Censorship is generally at odds with democracy, which is characterized by the institutionalization of civil and political rights, civil liberties, and the existence of an opposition. Although free expression is a central democratic value, some opponents of censorship do recognize exceptions. One of the most commonly proposed exceptions is of child pornography, due to its connection with individuals or organizations that commit crimes such as pedophilia. Supporters of censorship in such cases argue that the right to free expression is outweighed by the need to prevent the extraordinarily harmful impact on children caused by the creation and distribution of child pornography. In contrast, critics of censorship with respect to child pornography argue that the connection between pornography and actual harm to children has not been established. Although some proponents of free speech argue that it has precedence over the defense of children's interests, most Western democracies ban the possession, production, and distribution of child pornography.

Hate speech is another controversial subject of censorship; again, supporters argue that hate-filled speech encourages actual violence and that this potential harm outweighs the value of free expression. For example, such arguments are made to justify the censorship of neo-Nazi propaganda both in the United States and in Europe.

The use of censorship may be seen not only in a formal way—that is, through officials entitled to control political information using known political means—but also through informal mechanisms, such as threatening independent journalists, punishing a critical press through the denial of paid advertisements, and so forth. These informal censorship mechanisms can be found in young democracies established after military or authoritarian regimes, in which there are political constraints inherited from the prior political order limiting the institutionalization of political rights or caused by the lack of democratic beliefs among the governing elite. Referred to by Wolfgang Merkel as "defective democracies," they may simply not have repealed the previous regime's restrictions on press freedom or may consider that censorship is necessary in certain cases for reasons of national security. In this latter case, it is used to restrict publication of information about issues that are considered very sensitive for the armed forces and to avoid possible tensions and ensure the consolidation of democracy. In defective democracies or hybrid regimes, censorship by governing officials is exercised through formal and informal mechanisms, including threatening journalists. The elimination of censorship is one of the priority demands of an agenda that seeks to establish real democracy. In delegative democracies, such as Venezuela during Hugo Chavez's presidency (in power since 1999), public officials use a variety of methods to limit freedom of press, including not only censorship but also actions such as the cancellation of TV and radio licenses. Although censorship is formally rejected in established democracies, freedom of speech can be restricted not as a governing decision but as a consequence of the ownership structure of the media, when business people seek to influence public opinion with their own beliefs or political views, rejecting the dissemination of opposing views. In these situations, editors may employ extensive self-censorship to block the dissemination of news and ideas that may damage the government.

Censorship can be imposed *preventively*—by defining conditions for the dissemination and circulation of ideas through any channel—or ex post facto, with the authority sanctioning those who infringe the norms that establish limits on freedom of speech.

Although censorship is incompatible with democracy, this does not mean that public officials lack the resources with which to attempt to persuade the media to block news and information that may be damaging or favor reports that may be negative to the opposition. One of these key tools is the placement of advertising, which can be an important source of media revenues.

With the development of modern technology, particularly the expansion of Internet, the imposition of censorship is more difficult than ever, posing enormous problems for public officials seeking to limit the dissemination of information and news critical of the government. The Internet opened a new channel of communication for dissidents in nondemocratic regimes (e.g., Cuba and China) to denounce abuses and make their demand for pluralism and freedom known to the world. These actions, in turn, trigger support from NGOs and political parties and governments in democratic countries, which helps protect the authors from reprisals. Large international corporations—notably Google—have also emerged as new players in favor of greater pluralism in nondemocratic regimes or defective democracies.

Censorship plays a very important role in authoritarian regimes. It serves as a tool to ensure that limited pluralism—a central element in their identity—remains within the bounds established by the authorities. These regimes allow certain sectors of the elite access to, and even control of, the media and publishing houses through which they seek to gain political power or influence while also restricting the access previously enjoyed by other sectors of the elite. Self-censorship is widely applied in these regimes by editors of private media companies that support the regime, are interested in its continuity, and are particularly strict in preventing the publication of negative news, particularly that referring to opposition activities or conflicts and divisions within the governing elite.

Francisco Franco's Spain (1939–1975) provides an example of the evolution of censorship, which was not applied uniformly during his regime but underwent a number of changes as a result of conflicts among factions of the ruling elite, which had more influence when the regime was consolidated. With the consolidation of political power and in a context of economic development and social progress, there was a tendency to relax censorship in order to allow the press to play a role in controlling excesses, abuses of power, and, particularly, corruption among top officials. This explained the press law of 1966 in Franco's Spain, which began a certain liberalization of the official press, bringing greater transparency about the decisions of the authorities. However, the information concerning abuse of power, particularly of economic character, was used by sectors of the governing elite to weaken the power position of other factions, provoking a scandal that led to a cabinet reshuffle and the dismissal of ministers involved in this conflict. However, it also permitted the appearance of opposition publications and even magazines (e.g., *Cuadernos para el Diálogo*), which had costs for the regime in terms of a significant increase in negative political information since they published news about cases of corruption and conflicts within the government and the elite. As a result of these direct and indirect effects of the liberalization of the press, censorship was, in practice, reestablished.

In the context of liberalization of authoritarian regimes, the relaxation of censorship and self-censorship was widely extended, as in the Pinochet regime in Chile (1973–1990) during a period of political instability caused by the economic crisis in 1983 to 1984. Independent radios and magazines could inform in a broader way on several political facts that had been censored earlier, including repression against opposition organizations. However, the government continued threatening and punishing journalists with formal and informal mechanisms and applying censorship in particular cases. The political space opened by the weakening of censorship was widely used by journalists and media, leading to a significant increase in political information, including adverse information, and favored the development of the opposition and, as a result, the weakening of the political bases of authoritarian rule.

Carlos Huneeus
Universidad de Chile
Santiago, Chile

See also Hybrid Regimes

Further Readings

Jones, D. (Ed.). (2001). *Censorship: A world encyclopedia*. London: Routledge.
Norris, P. (Ed.). (1997). *Politics and the press*. Boulder, CO: Lynne Rienner.
Norris, P. (2000). *A virtuous circle: Political communications in postindustrial societies*. Cambridge, UK: Cambridge University Press.

CENTRAL BANKS

Modern central banks are public, nonprofit, economic, and political institutions, with special functions derived from their command over monetary resources. They shape monetary policy, have a strong influence on exchange rates, and are the guardians of financial stability. Their decisions influence economic variables determining growth, output, and national development policies and have a bearing on the conditions of international financial and monetary cooperation. As a general rule, central banks are regulated by a mandate from government(s) specifying their goals and the conditions under which control over monetary resources is exercised.

After situating the topic in its theoretical context, this entry discusses the rules and social institutions on which central banks' command over monetary resources ultimately rest; stresses the establishment of central banks as a trustworthy system of monetary authority, at the domestic and international levels; and examines the way in which the institution's power has been recast and strengthened in the financial globalization era, in connection with the widespread trend toward central bank independence from elected governments in old and new democracies.

Central Banks as Substate Actors: Theory

An updated review of existing central banks shows a greater variance than is assumed in the dominant rational expectations theories framing the universalistic prescriptions and the role models recommended over the past 30 years. Influence is noticeable, however, in two global trends: (1) in terms of goals, the convergence toward the primacy of price stability, and (2) in terms of the mandated use of monetary resources, the dramatic increase in the number of institutions that acquired statutory independence vis-à-vis governments. Yet the actual variance in the practices and mandates of those institutions across countries and regions limits the scope for generalization. Moreover, central bank functions and powers have been recast over time since the establishment of the first institutions of this breed in the mid-19th century.

A central bank's status as a public, nonprofit institution is the outcome of an evolutionary process, based on the rules and social institutions that underpin its three core functions: (1) as banker to the government, (2) its monopoly of note issuance, and (3) as banker to the banking system—including that of last-resort lender. These functions refer to the unique relationships between the institution and its principal clients: the government and the profit-maximizing financial institutions, mainly commercial banks.

The Relationship Between the Central Banks and the Government

As a public national institution, the central bank derives its functions from a state-backed power as the sole issuer of a unique currency circulating in a geographical territory—to which it has been granted the status of legal tender (*cours forcé*).

Parallel to the processes of political delegation, a contractual, economic relationship with the government involves mutual economic interest and economic calculation. Lending to the government, at an agreed fixed interest rate, generates profits alongside those drawn from currency issuance. For the government, having "its own bank" ensures cheaper deficit financing; its power as principal borrower is used to press for a minimum rate of interest. On the other hand, central banks typically stand out as major holders of a government's debt, insofar as they invest their capital in government long-term debt instruments, mainly bonds. Where a domestic public debt market exists or is being established, voluntary loans from private savers provide an additional source of deficit financing. It is a market monitored by the central banks—adding the function of debt management to their other functions.

While the government's preference for cheap finance is an intrinsic component of its relationship to their central banks, deficit financing through monetary expansion—that is, through the monetization of the government's debt—was exceptional in industrialized countries. Long-term reviews of past practices supplied by Charles Goodhart's team and by Barry Eichengreen indicate economic as well as political reasons for central banks to be congenitally inflation and risk averse. Because they are major holders of the government's debt and because they already profit from lending to the government, they have an overriding interest in keeping the real value of its loans stable. There is also a reputation dimension to its practice, in that the strength of the currency issued is closely related to the creditworthiness of the issuer and to the commitment of the issuing institution to price stability. Prior to the 1970s, the commitment of central banks to convertibility and/or price stability was overridden in only one circumstance: when the survival of the state was threatened by wars. Thus, the "inflationary bias" of economic policy and the temptations of monetary financing of government spending in industrialized countries are late developments that call for a more contextualized analysis of the unstable monetary "regime" in the financial globalization era.

Prompted by two "structural" factors, some developing countries, mainly in Latin America, resorted to an "inflation tax." The first factor was a weak tax system, reflecting the limited capacity of the state to establish itself as a widely accepted fiscal authority. The second was that the inflation tax operated as a conflict-minimizing device to socialize the costs of "catching-up" industrialization. This was typically, as in the case of Brazil, where the state's commitment to a proactive strategy of rewarding the winners was based on the reiterated distribution of skewed fiscal and monetary incentives to selected economic agents.

Bank to the Banking System

As bankers to the banking system, central banks play a leading role mainly vis-à-vis commercial banks and less effectively vis-à-vis other profit-maximizing financial institutions. Such a role is, first of all, one of guidance and monitoring. Because commercial banks act simultaneously as deposit takers and loan makers, they are credit multipliers. Since credit is money, their individual microeconomic decisions, guided by profit-maximizing considerations, have a direct bearing on the increase or reduction in the stock of money available in the economy. Central banks' overriding concern with financial and economic stability gradually evolved toward a proactive role in averting this source of volatility. Their function as lender of last resort to the banking system—asserted always in the wake of crises—revolves around the bank's systemic responsibility, in exchange for enhanced powers of supervision and regulation of commercial banks' behavior. While the necessary function of lender of last resort is called for and is reiterated in crises, it is through its open market operations that the leadership of the central bank over profit-maximizing institutions is exercised. To maintain day-to-day control over the short-term nominal interests rates in all market conditions, the central bank purchases or sells financial securities for cash to change the monetary base. Its power to influence the money supply and other economic variables in this case is *exercised through the markets*. In other words, through a complex system of incentives and penalties—mainly the bank's discount rate—it may induce commercial banks to proceed with their lending strategies (or dissuade them). Because the central bank can impose the ultimate conditions for granting credit, it enjoys the "magical powers" of creating or destroying money. But it does so by relying on market mechanisms—namely, on its monopolistic position in the credit market and on the predictably self-interested response of profit-maximizing institutions to its interest rate policy.

Set against this background, the use of the term *central bank* is equivalent to and interchangeable with that of *monetary authority*—a use that is common to economists of all persuasions. This identity relies on three assumptions. One, wherever a market exists, financial firms and investors act independently from one another. Two, the central banks' power over monetary resources is that of a rule maker—an extension of the state's capacity of enforcement. Three, its creditworthiness derives from its governance capacity in terms of price and financial stability, which in turn derives mainly from its technical credentials.

Political scientists assume a political economy approach to the international monetary regimes in which central banks operate as rule based and contingent. They take as a premise that domestic and world markets are embedded in and permeated by social institutions. From this perspective, the history of the international monetary order (or disorder) is thought of as "the history of the construction and demolition of rules, constitutive and regulative, explicit or tacit, substantive and procedural," in the words of Bruce Hall (2008, p. 10). Central to this cognitive framing are the politico-economic processes underlying the establishment or disruption of a policy consensus around a monetary regime.

The International Political Economy and Central Banking

A major distinctive feature of the current monetary and financial regime is the rising power of central banks vis-à-vis established governments and the redefinition of their authority vis-à-vis cross-border market actors. To situate such processes requires a shift of focus from its role as a predominantly substate institution to one including its capacities as a public transnational actor in a multilevel system of global governance. The bank's command over monetary resources has been recast in tandem with new modes of interaction with government(s) and with market actors. Those shifts were carried out within the cognitive framing of the rational expectations revolution—centered on the notion of central banks' credibility with the financial markets. They are closely linked to a number of developments that distinguish the post–Bretton Woods monetary era, extending from the 1970s to the present, from its predecessors: the classical gold standard (1844–1945), the interwar decades, and the Bretton Woods system (1945–1971). While the previous systems constituted international monetary regimes proper, whose rules and social institutions were designed and enforced by the United Kingdom (UK) and the United States, respectively, the post–Bretton Woods is considered by many, including Barry Eichengreen, as a "nonsystem." Successive failures at establishing a durable international monetary order and a great deal of technical experimentation underlie the process by which the powers of the central banks were recast.

The issues around which the practice and theory of central banking came to revolve are explored by the monetary economist Stanley Fisher. On the one hand, the inflationary tendencies are brought about by the conflict between the short-term and the long-run effects of monetary expansion. On the other, it should revolve around the conflict between the need to shield central banks from the political pressures underlying the monetary financing of government spending and the principle of accountability to the public.

Central Banking and Governments

Central banks' current powers are tied to critical shifts in their relationship with their constituencies and in the cognitive maps adopted by central bankers. A salient aspect of those developments is a shift in the priorities of monetary management toward the primacy of price stability over other policy goals, such as the promotion of full employment and maximum output. The underlying politico-economic framework is shaped by the threat of runaway inflation in the dominant capitalist democracies, in the 1970s and 1980s, mainly in the United States and the UK, which was an unprecedented experience in peacetime. The international political consensus underlying the Bretton Woods regime had already been shaken in 1971 to 1973 by President Nixon's unilateral decision to detach the dollar (to which all other currencies were pegged) from the gold standard so that dollars could no longer be converted to gold. Because the dollar was the reserve currency, severe losses in the value of foreign reserves held in dollars were inflicted on the other members in the international monetary system. The ensuing legitimacy crisis was compounded by the refusal of Germany and France to import inflation engendered in the United States.

A matter of interest to the social scientist is brought to the fore by this episode. It highlighted the capacity of a powerful state, the hegemon, to deflect the costs of an overdue domestic adjustment to trading partners. It also highlights the critical role of transnational political coordination in counteracting that mode of power. The unique experience of Europe is a case in point: Both that episode and the crisis of the European Exchange Rate System in the early 1990s propelled the gradual crafting of the institutions that underpinned the Maastricht Agreement provisions aimed at establishing a regional monetary order. The powers of national central banks were redefined accordingly, in tandem with the transfer of their command over monetary resources to the European Central Bank. Insofar as it involved a long-term project, multiple acts of sovereign political delegation, and democratic deliberation, the European Central Bank is a unique case of establishing a transnational monetary authority through statecraft, in the sense explored by Lourdes Sola and Laurence Whitehead (2005) in connection with developing countries.

It is within this shifting monetary, financial, and political context that the new powers and functions assigned to central banks vis-à-vis the governments become meaningful. The centrality of the monetary policy and the monetary counterreaction to the dominance of Keynesian thought marked out a paradigmatic shift away from the postwar economic dirigisme structured around demand management and based on exchange rate, price, and interest rate controls. It brought to an end the subordination of central banks to domestic fiscal authorities.

The empowerment of such institutions, in particular the U.S. Federal Reserve System, highlights in what sense the exercise of monetary authority—as a special mode of political authority—can be strengthened and redesigned within the domestic framework of liberal constitutionalism. Three dimensions are of interest: (1) the strengthened role of central banks as rule makers, (2) the political dimension of their effective capacity for monetary governance, and (3) the relative power of the hegemon—the United States, as the issuer of the international reserve currency, the dollar—to lead the international arrangements that deflected the costs of adjustment to developing countries.

The authority of the Federal Reserve as a substate actor is constitutionally specified and limited. Its autonomy vis-à-vis the executive is constrained by the oversight of the U.S. Congress and, in terms of its goals, by its dual mandate—price stability and full employment. Neither these constraints nor the prospect of the huge social and political costs inseparable from recession and unemployment barred the option for a shock therapy to curb runaway inflation—basically a dramatic increase in the U.S. interest rates by Federal Reserve Chairman Paul Volcker. The political dimensions of the decision-making processes underlying a sharp turnaround of monetary policy remain understudied—despite the evidences that it was premised on successive acts of political delegation from both the executive and the legislative powers.

It is widely agreed that the Federal Reserve's capacity for monetary governance, in terms of its success in curbing inflationary expectations, is better explained by its political components than by its technical ones. Monetary economists, practitioners, and political economists agree that as a technical exercise, it was hardly a success. Volcker's economic objective was achieved thanks to his decision to emphasize price stability and to de-emphasize (also on theoretical grounds) the possibility that monetary policy could affect the level of output. This, in turn, relied on the flexible, dual-mandate design of the Fed and, above all, on Volcker's intellectual authority and moral credentials. To many practitioners, such as Alan Blinder, this is a critical asset to forge the near unanimity required in the collegiate decisions within the Fed. Also, insofar as monetary policy works through the markets, the technical and moral credibility of practitioners in the eyes of market actors are critical. The technical complexities of monetary management make this mode of personal, contingent authority a powerful instrument of persuasion and legitimation in the deliberative processes in Congress—and in the eyes of the relevant constituencies.

The systemic international functions of the Fed and other core central banks—both as rule makers and in their capacity of governance—relate to one of the most consequential implications of the dramatic rise in U.S. interest for developing countries, particularly for Latin America: soaring indebtedness, vulnerability to further external shocks, and the withdrawal of foreign investors for 9 years. Central banks in core countries were called to lead the coordination of the international arrangements aimed at deflecting to debtors the costs of the U.S. adjustments. The threat that their rising debts might feed on the international banking system prompted the emergence of new institutional arrangements, in concert with the International Monetary Fund (IMF) and the representatives of the private banking system. The responses to the "debt crisis" in Latin America were designed by the committees organized to this end and enforced on a country-by-country basis. They were conditioned on the implementation of structural economic reforms, including deregulation of the domestic financial systems and, ultimately, central bank independence. Competition for credit and foreign investments explains the compliance with most of those prescriptions. Sylvia Maxfield explains the move toward central bank independence in developing countries as driven by the need to obtain credibility in the eyes of increasingly competitive financial market actors. However, to the extent that this is a global trend across five continents, the recasting powers of the institution at the global level must be specified too.

Independence as Global Governance: The Emergence of Private Authority

Underlying the *problématique* of central bank independence is a critical shift in the mode of interaction between the "monetary authority" and market actors. The rising influence of cross-border financial transactions is closely linked to the empowerment of foreign exchange markets and disintermediated debt and bond markets in individual nation-states. The intensified competition among banks and between banks and other financial intermediaries in the past 3 decades continues to be central to this process. Competitive pressures explain also the proliferation of "financial innovations"—that is, the recurring creation of new debt/credit instruments intended to make them more attractive to borrowers and to lenders, both corporate and public. Such innovations lie outside the direct control of domestic central banks, are enhanced by information technology, and are driven by the deregulation and the liberalization of domestic financial markets.

In an international system where there is no world government, governance capacity is found in multiple sites of monetary (and fiscal) authority—public and private, national and transnational. Hall hypothesizes that central banks are integrated into an emerging new system of global multilevel monetary governance. Public transnational institutions endowed with supervisory and coordinating monetary powers include the Bank of International Settlements (BIS) and the European Central Bank. The fact that adherence to such rule-based systems is voluntary—and applied domestically by statutory authority—indicates that its implementation depends also on domestic political variables. At the national level, the executive is the main locus of monetary governance; a central bank may share them with fiscal authorities if it has not been granted independence by the government.

In this context, private market actors operating in the foreign exchange and disintermediated bond markets are empowered to adjudicate the credibility of fiscal and monetary policies through the operation of the markets. This function is complemented by private rating agencies that grade the market of sovereign, corporate, and municipal debt instruments. All such shifts reflect the unprecedented emergence of private authority conceptualized by Thomas Biersteker and Bruce Hall.

It is against this shifting context of deregulated, liberalized financial markets—where all money is fiat money—that the global trend toward central bank independence becomes meaningful. Constructing a trustworthy monetary authority is about making credible its long-term commitment to price stability, ensuring private and public wealth holders that their loans will be redeemed at their real value, plus a risk premium. Insofar as central banks can directly control only very short nominal rates of interest, the technical question of how monetary policies will work effectively through the markets is a matter of great concern. Economists and practitioners of all persuasions agree that it has been settled technically along the lines of the rational expectations theory. In other words, monetary policy can work through the markets when central banks properly anticipate and coordinate future expectations of economic agents. For this reason, the case for central bank independence at the global level has been strengthened by the rational expectations school and its analysis of the conditions necessary to attain credibility with financial markets. The central bank's independence from governments and transparency in communicating its long-term future monetary strategies are the major conditions enabling the institution to achieve credibility with market actors. In this intellectual and economic context, the economic and financial impacts of central bankers' utterances and communicating strategies can hardly be exaggerated.

While rationalists assume a self-equilibrating free market of actors acting independently and including in their calculations the expected impact of monetary policies, constructivists assume a relational mode of interaction between them and among market actors. It rests on shared understandings of the rules, norms, and values that underlie what are and what should be considered best practices—underpinned by epistemic communities of monetary economists and central bankers. Hall, a political economist of a constructivist cast, sets the generalized trend toward central bank independence in the context of multilevel global governance: The powers assigned to the institution are related to its ultimate role as the only nominal anchor to achieve international stability. Policy convergence, in this approach, is meant to drive the most critical structures down to the lowest possible national level of governance so that a uniformly effective impetus toward price stability can be provided. Whether or not this desideratum implied in the rational expectations theory is feasible is a quintessentially political problem. It is constrained by the current shifts in the axis of global monetary power, in particular the diversity of monetary strategies of new, "catching-up" countries. China's central bank is the most extreme case of subordination to the unique strategies of the government, aimed at national development and integration into the world market.

There are, however, many points of intersection between the two schools. Most important among them is the unspoken assumption underlying their prescriptions on how to reconcile two desiderata: (1) shielding the monetary authority from political pressures and (2) political legitimacy. Both schools take for granted the political framework within which central bankers operate in core democratic governments—that of liberal constitutionalism. The solution is then settled by means of the distinction between "full operational independence" and "goals independence." The discretion of central banks, in the first case, would be limited to the choice of monetary tools to achieve the goals ultimately set by elected politicians. Basically, these tools provide free access to periodically publicized reports on the fiscal practices, the economic models, the forecasts and outcomes of monetary policy, and the accountability of elected politicians, in periodic hearings of Congress.

Emerging market democracies, in contrast, confront the task of reconciling integration into global markets, with its ensuing fiscal and monetary disciplines, while building the institutions and policy consensus indispensable to legitimate the construction of a monetary authority. It involves a long-term project, an open-ended process of democratic deliberation, leadership, and the development of constituencies supportive of price and financial stability as priority goals. In this regard, they constitute experiments in statecrafting monetary (and fiscal) democratic authorities.

Lourdes Sola
International Political Science Association and University of São Paulo
São Paulo, Brazil

See also Administration; Bureaucracy; Democracy, Theories of; Democracy, Types of; Economic Policy

Further Readings

Best, J. (2005). *The limits of transparency: Ambiguity and the history of international finance.* Ithaca, NY: Cornell University Press.

Bierstaker, T., & Hall, B. R. (Eds.). (2002). *The emergence of private authority in global governance.* Cambridge, UK: Cambridge University Press.

Blinder, A. (1998). *Central banking in theory and practice.* Cambridge: MIT Press.
Blinder, A. (2004). *The quiet revolution: Central banking goes modern.* New Haven, CT: Yale University Press.
Capie, F., Goodhart, C., Fisher, S., & Schnadt, N. (Eds.). (1994). *The future of central banking: The tercentenary symposium of the Bank of England.* Cambridge, UK: Cambridge University Press.
Eichengreen, B. (2008). *Globalizing capital: A history of the international monetary system* (2nd ed.). Princeton, NJ: Princeton University Press.
Goodhart, C. (1988). *The evolution of central banks.* Cambridge: MIT Press.
Hall, R. B. (2008). *Central banking as global governance: Constructing financial credibility.* Cambridge, UK: Cambridge University Press.
Maxfield, S. (1997). *Gatekeepers of growth: The international political economy of central banking in developing countries.* Princeton, NJ: Princeton University Press.
Posen, A. (1993). Why CBI does not cause low inflation: There is no institutional fix for politics. In R. O'Brien (Ed.), *Finance and the international economy* (pp. 40–65). Oxford, UK: Oxford University Press.
Sola, L., & Whitehead, L. (Eds.). (2006). *Statecrafting monetary authority: Democracy and financial order in Brazil.* Oxford, UK: Centre for Brazilian Studies.
Whitehead, L. (2002). On monetary authority. In L. Whitehead (Ed.), *Democratization: Theory and experience* (pp. 136–164). Oxford, UK: Oxford University Press.

CHANGE, INSTITUTIONAL

Scholars in political science have approached the problem of institutional change in two very different ways that can be explained by the type of institutions on which they have focused. Institutions are most commonly understood as the *rules of the game*. As such, they can be written formal statutes, such as constitutions, electoral systems, and legal codes. They can also be generally established social norms, such as routines, customs, and habits. In almost all accounts of institutions, they are understood as relatively stable and persisting entities that do not change easily or instantly. The broad notion of institutions and the focus on continuity imply that understanding institutional change is both complicated and equivalent to understanding large-scale social and historical change.

The reasons why institutions are usually seen as durable and difficult to change vary. One is that institutions give rise to self-reinforcing "feedback" mechanisms that may or may not be purposively designed. For example, a social insurance program may be organized so that it covers a majority of the population, thus creating not only social protection but also support from the political majority. Or it may cater to broadly held norms about social justice in the population. In either case, a self-reinforcing mechanism is at work. Another reason for institutional stability is that none of the actors that are involved have an incentive to change. For example, political parties in a "first-past-the-post" two-party political system have little reason to change to a proportional electoral system, or vice versa. Third, formal institutions may become dominating social conventions, and agents may have difficulties imagining a different institutional order.

Explanations for institutional change can be divided into three broad categories. Change may occur as a result of unpredictable and unforeseen "exogenous shocks." For example, countries may radically change their institutions when hit by an international economic crisis or when drawn into a war. A second type is a change through an evolutionary functionalist logic. Institutions that best suit the underlying structural changes survive through the operation of some kind of selective mechanism, and institutions that do not follow this functionalist logic are weeded out by the competition from more successful institutional orders. Third, institutions may change by intentional design by strategically acting as agents that construct new institutions that serve their future interests. However, the outcome of such strategic design may not always be in line with the agents' intentions because outcomes from institutional changes are hard to predict, especially if there are simultaneous changes of other institutions outside the control of the designing agents.

The Two Schools of Institutional Change

There are currently two very different approaches to the problem of institutional change. In development research, including research in comparative political economy, there is now almost a consensus that the problem of massive poverty (and the many resultant social ills) in most developing countries is due to their dysfunctional social, legal, and political institutions (e.g., Daron Acemoglu & James Robinson, 2008). This problem, also known as the "good-governance" or the "quality-of-government" problem, has been well captured by the economist Dani Rodrik (2007, p. 153), who argued that the encounter between neoclassical economics and developing societies has revealed the great extent to which market economies rely on a complex and not well understood set of social, legal, and political institutions. Rodrik's list of such institutions is extensive: a well-specified system of property rights, effective regulation that hinders monopolies from dominating markets, uncorrupted governments, the rule of law, and social welfare systems that can accommodate risks. In addition, Rodrik adds the importance of informal institutions that foster social cohesion, social trust, and cooperation. Following Douglass North, Rodrik criticizes neoclassical economics for omitting the importance of such institutions and argues that the problem is that most economists usually take them for granted. In fact, most poor countries are plagued by a very different set of institutions, such as systemic corruption, systematic infringements on property and human rights, and low levels of social trust and lack of institutions that can handle individual and collective risks. In this type of analysis, the main problem is to explain why societies with a set of institutions that are clearly dysfunctional in terms of general social well-being generally seem unable to adopt the "new set" of institutions that Rodrik lists and that are known to lead to economic prosperity and social well-being. As Bo Rothstein (2005) notes, the central problem in this type of analysis is that the historical record clearly shows that dysfunctional institutions, such a systemic corruption, clientelism, and patronage, are very hard to change. Political and economic leaders in these systems may be able to change the institutions from above, but since they are the group that collects the most rents from the system, they have no incentive to induce change, at least not in a short-term perspective. Change of these systems from below is difficult because agents who want change face a formidable collective action problem. Thus, as Douglass North, John Wallis, and Barry Weingast (2009) have shown, in this research area, the main issue is the lack of institutional change despite the pressing human need for such change. For example, systemic corruption not only hurts economic growth, causes extreme poverty, and induces civil strife, but it also has devastating humanitarian consequences in areas such as provision of health care, level of infant mortality, and access to safe water.

However, in several recent studies of the rich welfare states and their systems of industrial relations and social provisions, the picture is very different. Here, change of important institutions occurs almost constantly in a mostly incremental process, either because agents are able to find compromises and viable solutions to common problems or as a result of power struggles. As James Mahoney and Kathleen Thelen (2009) and Thelen (2009) have shown, adaptation to new circumstances such as increased international competition for goods and services, globalization of financial markets, and new technologies in the production process may not always be easy in these societies and may be the outcome of difficult political disputes and difficult compromises. Still, the general picture is that in such societies, political and economic elites are, after some initial difficulties, usually able to gradually change central institutions to bring about the distribution of resources and power that they believe will follow from their ideas of new institutional designs.

A problem for our understanding of institutional change is that these two research approaches seldom communicate their theories, problems, or results. Why do the political and economic elites in some (rich, capitalist) societies usually manage to change institutions, while other (poor, developing) societies seem stuck with formal as well as informal institutions that are fundamentally detrimental to their need for economic growth and general social well-being? One reason for this apparent paradox may be that these different approaches to institutional change are dealing with institutions that are fundamentally different.

Two Logics of Institutional Change

The idea that institutions include not only formal but also informal rules implies that it is difficult to distinguish them from a society's basic cultural traits. In explaining institutional change, this becomes problematic because while it is possible to change codified systems of rules and "standard operating procedures" through methods such as direct and deliberate political intervention, this is much more difficult with institutions viewed as shared mental models and other such generally held basic beliefs that are rooted in a society's historically established culture. For example, one such informal institution is people's belief that others in their society are, in general, trustworthy. Comparative survey research shows that this informal institution varies to a large extent among countries. Clearly, changing such a system of beliefs is not the same thing as changing a piece of social insurance legislation or a system of wage negotiations. Thus, as a first distinction, we can differentiate between two basic logics of institutional change, namely, change in a society's formal institutions and change in its informal ones.

A second distinction has been suggested by George Tsebelis (1990)—namely, between "redistributive" and "efficient" institutions. *Redistributive institutions* are rules that move resources or power from one group of agents to another. A familiar example would be most tax systems and systems of industrial relations. Historically, one can also think of slavery as a formal redistributive institution. Examples of redistributive informal institutions would be professional networks that give status and prestige to a particular set of its members. Another example is economic systems in which business contracts and career systems follow family lines. Third, many political systems are characterized by informal clan-based nepotism or various forms of corruption.

Socially efficient institutions, on the other hand, have quite the opposite character since their purpose is to improve the welfare of all actors in a system of exchange. As such, they are genuine public goods, implying that they serve a common interest. According to Elinor Ostrom (1990), seen in light of noncooperative game theory, these are institutions that make it possible to avoid situations known as suboptimal outcomes in prisoner's dilemma, collective action, or tragedy of the commons types of interaction. Rothstein (2005) notes that in the closely related theory about social dilemmas, socially efficient institutions make it possible for agents to avoid ending up in situations known as social traps. These are all situations in which the agents know that they would all be better off if they could collaborate to establish a common set of institutions; however, because they do not trust that the other agents will adhere to the rules or contribute to the costs of establishing and running the institution, the individual agents have no incentive to collaborate. This lack of trust leads to a situation in which all agents are worse off. For example, a corrupt doctor in a public health system may realize that all would gain by ending corruption, but it becomes meaningless for the individual doctor to stop taking bribes if she or he cannot trust that most other doctors would do the same. Similarly, it makes little sense for the single judge or police officer to stop taking bribes if he cannot trust that (almost) all his colleagues will also refrain from this practice. In an emerging democracy, it makes little sense to be the only political party that does not tamper with ballot boxes. The list of examples illustrating this problem is endless.

Examples of socially efficient formal institutions are the rule of law, protection of property rights, and an honest and impartial civil service. In economic terms, such institutions are public goods because they dramatically lower what are known as transaction costs and thereby contribute greatly to economic prosperity. Universal social insurance systems can also be categorized as socially efficient institutions since they handle individual risks in a more cost-efficient way than do market-based social insurance systems.

The most well-known example of informal socially efficient institutions is the existence of generalized trust that is widespread in the population. Institutions such as this increase the likelihood that other agents will not use opportunistic or treacherous strategies but instead will behave in a reciprocal manner. Thereby, generalized trust also decreases transaction costs, but in a less costly way than if agents were to use only the formal legal system.

The reason why it is so difficult to change from dysfunctional to socially efficient institutions (e.g., establishing honesty in a corrupt civil service or changing political exchange from clientelism to impartiality) is that such a change necessitates the trust among agents that (almost) all of them will play by the new rules from time t_1 and that such a trust is not easily manufactured by political means. As Ostrom (1990) has argued, socially efficient institutions can be seen as collective action problems of the second order.

It should be added that the distinction between socially efficient and redistributive institutions is a theoretical construct and that many, if not most, institutions have traits of both redistribution and social efficiency. Why the distinction is useful may be better understood from the following parallel example. Most political scientists are willing to distinguish between democratic and nondemocratic political systems. However, all democratic systems contain nondemocratic elements, and most nondemocracies feature some kind of system for popular influence and for establishing consent.

The reason why scholars studying the rich democracies and scholars studying developing countries have come to such different conclusions about the possibility for institutional change is that they study fundamentally different types of institutions. The former usually concentrate on redistributive institutions for which the analysis of power constellations explains change. The latter study the lack of socially efficient institutions, which are difficult to establish because they are public goods. According to Acemoglu and Robinson (2008), the problem is that there is currently no generally accepted theory for how they can be established or why they are reproduced. Following standard assumptions about self-interested rational agency, socially efficient institutions of the type described above would not exist, and poor countries that lack them would not be able to change from inefficient to efficient institutions.

Bo Rothstein
University of Göteburg
Göteburg, Sweden

See also Development, Political; Institutionalization; Institutional Theory; Institutions and Institutionalism

Further Readings

Acemoglu, D., & Robinson, J. (2008). *The role of institutions in growth and development.* Washington, DC: World Bank.

Mahoney, J., & Thelen, K. (2009). *Explaining institutional change: Ambiguity, agency, and power in historical institutionalism.* New York: Cambridge University Press.

North, D. C., Wallis, J. J., & Weingast, B. R. (2009). *Violence and social orders: A conceptual framework for interpreting recorded human history.* Cambridge, UK: Cambridge University Press.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action.* New York: Cambridge University Press.

Rodrik, D. (2007). *One economics, many recipes: Globalization, institutions and economic growth.* Princeton, NJ: Princeton University Press.

Rothstein, B. (2005). *Social traps and the problem of trust.* Cambridge, UK: Cambridge University Press.

Thelen, K. (2004). *How institutions evolve: The political economy of skills in Germany, Britain, the United States, and Japan.* Cambridge, UK: Cambridge University Press.

Tsebelis, G. (1990). *Nested games: Rational choice in a comparative perspective.* New York: Cambridge University Press.

Charisma

Most basically, charisma refers to a rare trait found in certain humans that combines unusual charm and an ascribed "magnetic" quality of personality and/or appearance. Deriving from the Greek word χάρισμα (chárisma), it includes the meaning of "gift" or "divine favor," in particular with reference to innate and sophisticated personal communication and persuasiveness.

Charisma as a concept has been employed in several contexts, as a religious concept, a sociological concept, or a psychological concept, and also in journalistic and common language usage. Although impossible to operationalize accurately, charisma is often used to describe, or rather to label, a personality trait that includes the seemingly "supernatural" ability of some rare persons to lead, charm, persuade, and inspire others due to their "magnetic," alluring quality, without using power or threat. It refers to an ascribed quality, not to be achieved purposefully, in certain people who draw the attention and admiration (but also hatred if the attribution of charisma is perceived to be dangerous) of others.

As a political concept, charisma became famous most of all due to the analytical formulation by Max Weber (1864–1920) as one of his three types of legitimate rule (*legitime Herrschaft*). In addition to the rational-legal (-bureaucratic) rule (*legale Herrschaft*) and the traditional rule (*traditionale Herrschaft*), the charismatic rule (*charismatische Herrschaft*) was applied by Weber (1924/1947) to historical forms of political domination "resting on devotion to the exceptional sanctity, heroism or exemplary character of an individual person, and of the normative patterns or order revealed or ordained by him" (p. 215). In his writings about charismatic rule, Weber (1946) applies the term *charisma* to

> a certain quality of an individual personality by virtue of which he is considered extraordinary and treated as endowed with supernatural, superhuman, or at least specifically exceptional powers or qualities. These are not accessible to the ordinary person, but are regarded to be of divine or similar supernatural origin, and on the basis of which an individual is perceived and accepted as a "leader." . . . How the quality in question is ultimately judged from any more general ethical, aesthetic, or other point of view is, of course, irrelevant for purposes of definition. What only matters is how the individual is actually regarded by his "adherents" or "disciples." (p. 329)

Charisma thus denotes a relationship rather than an individual personality attribute. Should the strength of this belief fade, the domination of the charismatic leader can also fade quickly, which is why this form of rule can be very unstable, as the exercise of charisma is founded purely on its recognition by the leader's followers. Although charismatic domination evolves in contexts of traditional and/or rational-legal rule, it tends to challenge these forms of rule and is therefore interpreted by Weber as revolutionary. Although charismatic domination represents personal and noninstitutionalized leadership, Weber points to some processes of longer duration for which he employed the term *routinization of charisma* (*Veralltäglichung des Charisma*). Under the constant challenge of traditional and/or rational-legal forms of domination, he recognizes the partial institutionalization of charisma through the establishment of specified positions open exclusively to persons who demonstrate special personal qualities. Institutionalized charisma is also represented by *charisma of office* (*Amtscharisma*), which pertains to beliefs that certain officeholders, by virtue of occupying an office (e.g., priesthood), acquire certain special powers or qualities. By contrast, the pure personal charisma attributed to revolutionaries, prophets, and sages resists institutional influences; it is antithetical to stable authority based on fixed codes and customs.

Weber saw the charismatic leader as a vital agent of social, political, and religious change, which led to some criticism attributing an idealistic "Great Men" theory of history to him. In any case, charismatic leaders generally are said to arise in unsettled times permeated by disorienting sociocultural change. In such periods, unconventional political and/or religious groups arise composed of people who are fearful of the future, who hope that by placing their faith in some charismatic leader they will eradicate the past and protect their lives against unknown and unseen dangers. For example, the two most charismatic leaders in the 20th century are Mahatma Gandhi in India and Nelson Mandela in South Africa. Ayatollah Khomeini of Iran is an example of a leader with combined personal and office charisma.

The volatility of charismatic authority and of groups manifesting charismatic leadership has been a persistent theme. In essence, charismatic domination is unstable because it lacks both institutional restraints and institutional supports. The absence of institutional restraints on charismatic leaders coincides with the lack of institutional supports available to sustain a leader's authority. Charismatic domination is fundamentally a precarious state of affairs because the leader's claims to domination rest purely on subjective factors, as the perceptions of the followers of his or her extraordinary qualities may be situation specific and ephemeral. The charismatic leader must continually face the prospect that her or his special "gift" of "grace" will no longer be perceived the way it once was and that her or his authority will fade, as happened in the case of Kwame Nkrumah of Ghana. Charismatic leaders must continually be on the alert for threats to their authority from outsiders, dissidents, and rivals within the movement as well as from their administrative staff of traditional and/or rational-legal domination. The latter is generally oriented toward expanding the scope of its authority and rationalizing administrative procedures to the detriment of the charismatic leader's freedom of action, sometimes leading to his or her actual deposition.

An additional consequence of the lack of institutional supports for charismatic leadership involves the absence of regularized procedures for the transfer of authority, that is, the problem of succession. The failure to effectively institutionalize the charisma of the founding leader usually leads to intensifying factionalism and ultimately to the end of charismatic domination.

The precariousness and instability of charismatic leadership and its consequences in terms of group volatility, factionalism, and possible violent episodes represent instances in which the concept of charisma can facilitate the explanation of sociohistorical events. In this context, the concept of charisma has been criticized as basically a descriptive concept that labels rather than explains the power of leaders and the submission of believers. It has been suggested, further, that the development of controversial new movements and noninstitutionalized forms of charismatic behavior has been enhancing the importance of charisma. Yet charismatic authority and its concomitants in terms of the tendency to view social relationships and organizations in personal terms and to envision a messianic termination of present evils is generally thought to be associated more with traditional rather than complex modern societies.

Dirk Kaesler
Philipps-Universität Marburg
Marburg, Germany

See also Caudillismo; Legitimacy; Populism; Social Movements

Further Readings

Adair-Toteff, C. (2005). Max Weber's charisma. *Journal of Classical Sociology, 5*(2), 189–204.

Barker, E. (1982). *New religious movements.* New York: Edwin Mellon Press.

Neitz, M. J. (1987). *Charisma and Christianity.* New Brunswick, NJ: Transaction Publishers.

Weber, M. (1946). *From Max Weber: Essays in sociology.* New York: Oxford University Press.

Weber, M. (1947). *The theory of social and economic organization.* New York: Free Press. (Original work published 1924)

Weber, M. (1978). *Economy and society: An outline of interpretative sociology* (2 vols.). Berkeley: University of California Press.

Worsley, P. M. (1970). *The trumpet shall sound.* London: Paladin.

Christian Democratic Parties

Political families can be identified, following Maurice Duverger's approach of *familles spirituelles*, as parties that share the same ideological inspiration, by adopting Stein Rokkan's structural approach, which identifies parties according to their roots in social cleavages or in terms of institutional international networks, which connect ideologically similar parties in various countries. In the case of the Christian Democratic parties (hereafter CDPs), all these approaches provide fruitful hints in assessing the profile of this political family across countries, since the evolution of the ideological and social orientations coincided with changes in the international organization. This entry discusses some specific features of CDPs, their origins and their evolution over time, and their role in the present world.

CDP is a specific label that implies a rather narrow focus—parties sharing the "Christian-democratic ideology." However, the acronym *CDP* is often used more loosely. If we were to follow the structuralist and ideological approaches very strictly, all Christian parties would be part of this political family. To classify CDPs, Rokkan's framework can be used to refer specifically to religious parties that either arose in Catholic countries as the result of a church–state conflict or were formed when independence was sought by Catholics versus a dominant power as in Belgium, Ireland, and also partly Poland at the time of national independence after World War I. Because the formation of CDPs involved primarily the relationship between the Catholic Church (hereafter the Church) and the political sphere, Protestant denominations of Christianity play a minimal role in CDPs (with the partial exception of the Netherlands) and, thus, are not discussed in this entry.

However, to provide a complete picture, other Christian parties do exist in Europe—for example, in Norway since the end of World War II and in Sweden and Finland since the 1960s and the 1970s, respectively. But these parties represent only a very small number of voters. After 1989, some CDPs emerged in the new democracies in Central-Eastern Europe—for example, the Christian Democratic People's Party (Kereszténydemokrata Néppárt, KDNP) in Hungary, the Slovenian Christian Democrats (Slovenski krščanski demokrati, SKD) in Slovenia, the Lithuanian Christian Democratic Party (Lietuvos krikščionys demokratai partija, LKDP, which in 2001 merged with the Christian Democratic Union to form the Lithuanian Christian Democrats—Lietuvos krikščionys demokratai, LKD) in Lithuania, the Christian Democractic Movement (Kresťanskodemokratické hnutie, KDH) in Slovakia, and the Christian and Democratic Union/Czech People's Party (Křesťanská a demokratická unie/Československá strana lidová, KDU-ČSL) in the Czech Republic. All these parties and the various subsequent mergers never played a pivotal role comparable with their Western counterparts. In most of these cases, they did not even attain 10% of the votes. Most striking is the absence of a true CDP in the fervently Catholic Poland, except for a very short period.

Further, CDPs are present, in descending order of political influence, in Latin America, Africa, and Asia. In Latin America, CDPs played a significant role in various countries thanks to prominent leaders such as Eduardo Frei in Chile and Rafael Caldera in Venezuela. In 1949, these parties created a regional organization, the Christian Democratic Organization of America (CDOA). The formation of CDPs in Latin America followed a different route from the one taken in Europe. Some Catholic parties were present in Latin America in the 19th century, but they were much more conservative than the nascent Catholic political organizations in Europe at the time. After the 1930s, and especially after the 1950s, new, more genuine "Christian Democratic" parties emerged in Latin America, often in opposition to preexisting Catholic parties, as in the case of Ecuador and Uruguay. The most successful case among these parties was the Chilean Christian Democratic Party (Partido Demócrata Cristiano, PDC), which played a dominant role in the democratic periods before and after the military dictatorship of Augusto Pinochet. Other relevant parties are the Social Christian Party (Partido Social Cristiano de Venezuela, COPEI) in Venezuela, the Guatemalan Christian Democracy (Democracia Cristiana Guatemalteca, DCG) in Guatemala, and the PDC in El Salvador, all of which dramatically declined in the late 1990s, and the Social Christian Unity Party (Partido de Unidad Socialcristiana, PUSC) in Costa Rica and the union of the Christian Democratic Party with the Popular Democrat legislative group (Partido Demócrata Cristiano/Democracia Popular, PDC–DP) in Ecuador.

Even in their initial formation, CDPs in Latin America had a much looser relationship with the Church than did their European counterparts. However, they have caught up with European trends both in their emphasis on social concerns—at least for most of them—in the first postwar decades and then in relaxing their religious connection following the European trend of a more secular-conservative approach. Given the instability of the Latin American regimes and the frequent disruption of democracy (with the exception of Costa Rica), this analysis will focus mainly on the European context.

In Europe, a different approach to the study of CDPs, such as the "actor-based" one offered by Stathis Kalyvas, leads to the same focus on Catholic parties only (again with the exception of the Netherlands). According to Kalyvas's landmark contribution (in his book *The Rise of Christian Democracy in Europe*) "core European" CDPs were the by-product of Catholic political entrepreneurs' reactions to the mounting secular and anticlerical attitude of the liberal elites at the end of the 19th century. The Church did not sponsor the formation of religious/Catholic political organizations; on the contrary, it acted to limit, discourage, or even suppress such initiatives. This interpretation stresses the *autonomous* role of political entrepreneurs during the formation of the religious parties in Western Europe. The Catholic parties of the late 19th and early 20th centuries grounded their legitimacy in themselves; they did not depend on an external institution for their legitimation. Such internal legitimacy is an important asset for the institutionalization and organizational development of a political party. But despite these favorable internal conditions, the Catholic parties either collapsed or underwent profound modifications after World War II. However, even though they played a rather limited role in the prewar party systems, their role became pivotal and in some cases dominant in postwar Europe. Scholars disagree on the question of whether this shift in the status of European CDPs is best understood as a complete break between prewar and postwar parties or whether there was some continuity despite the change. This academic dispute can be better understood in light of a series of factors: (a) the attitude of the Church vis-à-vis the "modern world," and of CDPs in particular; (b) the ideological renewal of CDPs; (c) the novel postwar international context; and (d) the role played by CDPs in the various party systems.

Christian Democratic Parties' Formation and Evolution

The Church and the Modern World

As the first Catholic parties in Europe emerged at the turn of the 20th century, they were in more or less open conflict with the Church over the acceptance of the principles of liberal democracy. The Church still indulged in its self-representation of a "*societas perfecta*" with the Pope at its apex—as a sovereign among sovereigns—and maintained its condemnation of modern society and liberal-democratic principles, which it held were leading to a general and harmful "social disorder." The anticapitalist "social doctrine" of the Church as expressed in the encyclical *Rerum novarum* (1891) derives from this vision of society. However, in the same period, Catholic intellectual circles, especially in Italy and France, elaborated on the mundane ideology of "Christian democracy," which advocated the protection and promotion of Catholic interests through social networks and communities and the fostering of these positions in the political arena as well. In other words, Christian democracy recognized the opportunity offered by democracy for developing social activities based on Catholic principles and defending Catholic values (and interests) through democratic institutions. The Church radically condemned this ideology in the encyclical *Pascendi* (1907), declaring that the Christian-democratic ideology would cause "confusion between evangelization and the promotion of democracy." In accordance with this, the pontificate of Pius XI (1922–1939) aimed at restoring the "kingdom of Christ" against the liberal society.

This intransigent position was confronted during World War I and its aftermath. The war involved, in an unprecedented way, the religious hierarchies and the religious constituency in support of each respective state, and this relaxed the negative relationship between the Church and the liberal state. For the first time, the Catholic world overtly supported the state, ending that sense of extraneousness that had been growing since the 1880s. Moreover, the introduction of mass male suffrage enfranchised millions of new voters, and these voters needed a guide to prevent their becoming adherents of atheist socialism. As a consequence, the Church somewhat relaxed its negative attitude regarding political activities by Catholics and Catholic parties, even if it avoided any direct sponsorship.

CDPs did not exploit this "autonomy" from the Church by building up a solid identity and institutionalizing the party organization; on the contrary, they suffered from this distance and remained prone to adapting themselves to whatever position the Church took. Lacking open legitimation by the Church, they were constantly seeking the Church's approval of their policies. When the Church signed concordats with fascist Italy and Nazi Germany in the 1930s, the CDPs, with few exceptions, accommodated themselves to fascism and in some instances even supported it. The most dramatic case was provided by the Spanish Civil War (1936–1939). The radical anticlerical attitude of the Spanish Republican forces led the Church to sponsor, even with enthusiasm, General Francisco Franco's restoration and the intervention of the Italian fascist troops. The Church espoused the fascist regimes so easily because of its long-drawn-out opposition to modern society, whose liberal institutions and worldview the Church still despised. In addition, the economic depression was seen as a failure of capitalism, increasing the dissatisfaction with liberal society. The economic crisis reinvigorated Christian social doctrine, which was seen now as a "humanistic" third way between capitalism and communism.

The Emergence of Personalism

In the same period, however, a different elaboration was emerging, especially in Italian and

French intellectual circles. That milieu, where Emmanuel Mounier and Jacques Maritain played a crucial role, produced an original and quite powerful "ideology" of Christian inspiration: personalism. In a nutshell, personalism emphasizes the role of the person in society: The person is not conceptualized as the individual of the liberal tradition, an "isolated monad," but rather as a part of a collective body, intertwined with all the others through a network of associations and communities. Each person has intrinsic value *and* value as a member of a community or an intermediary body: Persons cannot be disconnected from their social environment, starting with the family. This new ideology played a dramatic role in moving CDPs (and to some extent, the Church) toward acceptance of the modern world and in distancing them from fascism. The most visible confirmation of this change came with the Pope's Christmas address in 1944, when for the first time he acknowledged democracy explicitly and clearly (*apertis verbis*). According to some scholars, however, Pius XII (1939–1958) accepted democracy only instrumentally, as the antagonist of communism, rather than for its own sake. Thus, according to Martin Conway and Tom Buchanan, Catholicism in the 1940s and 1950s remained intransigent and made few concessions to the modern world.

Postwar Christian Democratic Parties

At any rate, the CDPs that reemerged or were created *ex novo* after World War II proved much stronger than those of the prewar years. The international context, the Church strategy and profile, the CDPs' ideological-religious references, the party systems' new configurations, and the issue at stake are the factors that explain this change.

Church Profile and Strategy

World War II changed the image and the mass perception of the Church. Contrary to the antimodern, traditionalist, and mildly profascist stance of the prewar Church, after the conflict it emerged with a new image thanks to its charitable, comprehensive, and universal messages. France was an exception because of the deep involvement of the episcopate with the Vichy regime, which was partially counterbalanced by the support of the lower clergy for the anti-Nazi Maquis. The Church's acceptance of democracy and its more or less unequivocal accommodation with the modern world facilitated a positive reception of its role in a larger sector of the population than was the case previously. Even more important was the Church's modified attitude vis-à-vis the political presence of Catholics. Whereas in the interwar years the Church had kept its distance from the Christian parties and even despised them, during the war, the Church began to support CDPs and even sponsored their resurgence and the creation of new CDPs. This positive relationship provided the CDPs with the support of the Church organization: Priests became quite active, and parishes served as local branches or centers of support for the CDPs.

Anticommunism

One reason for this change of attitude lay in the new international setting. With the descending of the Iron Curtain and with atheist communism gaining power in Eastern Europe and making strides in Western Europe, too, the Church called on all religious persons to defend Christian civilization. And the CDPs followed suit. Anticommunism was the strongest and most successful appeal for the CDPs. They became the bastion against communist subversion and its antireligious politics. This appeal led to an enlargement of the electoral constituency; not only the faithful but the moderate voters as well were mobilized by anticommunism.

Location of CDPs in the Party System

CDPs gained a pivotal position in the various party systems. They were almost always a part of coalition governments in Belgium, the Netherlands, Luxembourg, Italy, and France until 1958 (and afterward the residual component in the United Democratic Front [UDF]). In Austria and Germany, however, they had to be in the opposition during some periods (the Austrian People's Party [Österreichische Volkspartei, ÖVP] in the 1970s and 1980s and the Christian Democratic Union/Christian Social Union [Christlich-Demokratische Union/Christlich-Soziale Union, CDU–CSU] in the 1970s and in the 1998–2005 legislatures in Germany). The CDPs' central position and broader ideological appeal were supported by the postwar collapse of the moderate-conservative bourgeois parties. These parties were discredited by their inability to counteract the totalitarianism of the interwar period or by their acquiescence to it. Where they attempted a resurgence after the war, it turned out to be a fiasco. The CDPs thus absorbed the conservative constituency too. In this way, the CDPs went beyond their traditional rural constituency, penetrating into the urban bourgeoisie (and sometimest the urban proletariat, too, thanks to the espousal of social welfare in the first years after the war, following the example of the French Popular Republican Movement [Mouvement Républicain Populaire, MRP], the most leftist and welfare oriented of all CDPs).

Ideology and Policy

The postwar CDPs' ideological bases were their reference to the Church on moral issues; the elevation of personalism as an encompassing "ideology"; the adoption of welfare systems combined with the principle of subsidiarity, placing the family at the center; an adequate dose of state intervention congruent with the social market economy; and a consensual/consociational conception of democracy, multilateralism, and peaceful coexistence. In particular, in the 1950s and 1960s, CDPs competed with communists and socialists on their own ground, presenting a social and "humanist" alternative to the Left's socialistic project. Later, the CDPs fully endorsed the "social market economy," which became their distinctive economic policy. All these elements have not changed up to the present time, except that state intervention has been downsized due to the neoliberal mood of the 1980s and 1990s. Nevertheless, CDPs were affected by the combination of the processes of modernization and secularization, which reduced the social basis of religion and, by the Second Vatican Council, had distanced many believers from active political support of CDPs. The Council had a tremendous impact on the Church and on most fervent Catholics. As far as politics was concerned, the Council fully recognized the principle of liberal democracy as such—and not instrumentally as a bastion against communism—and it disengaged the Church from day-to-day politics and from open support of CDPs. These changes provoked some criticism and backlash against many CDPs but not all of them. As far as their electoral fortunes were concerned, in the Netherlands and Belgium, the impact was strong, negative, and immediate; in Austria and Switzerland, perceptible but limited; and surprisingly quite limited in Italy. Only the CDU–CSU in Germany, thanks also to its interreligious nature, did not suffer any reversal.

The process of secularization and the decline or collapse of most "Christian Democratic" parties in the core of Europe (the Christian Democrats [Democrazia Cristiana, DC] in Italy, the Christian People's Party/Social Christian Party [Christelijke Volkspartij/Parti Social Chrétien, CVP–PSC] in Belgium, and the Christian Democratic Appeal [Christen-Democratisch Appè, CDA] in the Netherlands) favored a shift toward a more moderate and less religious orientation, following the imprint of the CDU–CSU. The evolution of the European Peoples Party (EPP) in the European Union (EU) shows this tendency quite clearly. EPP was progressively enlarged by moderate secular parties, which relaxed the religious tone while keeping some basic tenets of the CDP ideology: (a) the humanist personalism, (b) the centrality of the family, (c) the welfare concern and the social market economy, (d) subsidiarity, (e) the pro-EU attitude, and (f) the multilateral orientation in the international sphere.

Electoral Evolution

In the first elections after World War II, CDPs emerged as the strongest parties in France, Italy, Belgium, the Netherlands, Luxembourg, Austria, and Germany and as the third strongest in Switzerland. With the exception of France, where the MRP rapidly declined facing a newly established competitor, the Gaullist Party, and then collapsed with the advent of the Fifth Republic, the other CDPs retained their pivotal position for at least 3 decades. In the 1970s, their electoral fortunes and key positions in coalition governments declined. The Belgian and Dutch CDPs suffered the worst blows, losing votes and their dominant position up to the point where the three main Dutch religious parties had to create an interconfessional new party (the CDA) in order to retain their relevance in the political system. In Austria, the ÖVP lost its primacy to the Socialist Party in the 1970s and never recovered afterward. In Germany, the CDU–CSU, too, scored below the Social Democratic Party (Sozialdemokratische Partei Deutschlands, SPD) in the 1970s; it recovered in the 1980s but lost again in the 1998 elections. The most dramatic decline was experienced by Italy's Christian Democrats (Democrazia Cristiana, DC), a party that had always been in government since 1945 as the first national party; it was reduced to nearly 10% in 1994, losing almost 20 percentage points in 2 years (from 1992 to 1994). From the DC, a variety of parties emerged: The most relevant were the Italian Popular Party (Partito Popolare Italiano, PPI), its more direct heir, and the more moderate Christian Democratic Union (Cristiani Democratici Uniti, CDU) and Christian Democratic Center (Centro Cristiano Democratico, CCD), which then merged into the Union of the Center (Unione dei Democratici Cristiani e di Centro, UDC).

The crisis of many CDPs was due mainly to the process of secularization, on the one hand, and the reemergence of moderate-conservative parties, on the other, which in the first postwar years had left the political space empty. This reemergence brought a new ideological reference—neoliberalism and neoconservatism—which contributed to the marginalization of the two main elements of the CDP ideology, personalism and social concerns.

International Networks

The evolution of the international network of CDPs illustrates quite well the recent changes in this political family. Immediately after the war, in 1947, CDP leaders met informally and quite discreetly in the so-called Geneva Circle. That was an occasion for facilitating the relationship between the German Christian Democratic politicians and the other European partners, especially the French ones. It was meant mainly to promote mutual understanding among former rivals and to support the German chancellor Konrad Adenauer. Parallel to this initiative was the somewhat more formal Nouvelles Equipes Internationales (NEI), which was a gathering of Catholic parties and associations. The NEI worked together with the confessional parties of the European Coal and Steel Community (ECSC) and later in the European Economic Community (EEC) member states. In 1971, the NEI evolved into the European Union of Christian Democrats (EUDC).

This development of international networks in Europe was paralleled by a more global initiative to bring together all the CDPs in the world. Significantly, the Christian Democratic World Union (CDWU) was founded in Chile, in 1961, emphasizing in this way the relevance of Latin American parties and the truly international scope of the organization. The golden era of CDPs, as already stated, came to an end with the process of secularization, with the Church distancing itself from daily politics since the 1970s, and with the return with a vengeance of moderate-conservative ideology and political forces.

In the mid-1970s under the auspices of the CDU–CSU, a new organization was created, the European Democratic Union (EDU), which brought together CDPs such as the CDU–CSU and the ÖVP, and also secular moderate-conservative parties. This move influenced also the former EUDC, which evolved into the EPP in 1978. The new name, not by chance, did not include any reference to religious inspiration. The German CDU–CSU endeavor to enlarge the former CDP network to include secular moderate-conservative parties proved successful: The EPP expanded to become the strongest parliamentary group in the European Parliament in the 1990s. The tendency to loosen the religious connection was not counteracted by the core Catholic parties because of their growing marginality. Such trends affected also the international organization Christian Democratic World Union (CDWU), which in 1999 changed its name to Christian Democratic and People's Parties International.

Piero Ignazi
University of Bologna
Bologna, Italy

See also Christianity; Church–State Relationships; Parties; Party Systems; Religious Movements

Further Readings

Conway, M., & Buchanan, T. (Eds.). (1996). *Political Catholicism in Europe 1918–1965*. Oxford, UK: Oxford University Press.

Delwit, P. (Ed.). (2003). *Démocraties chrétiennes et conservatismes en Europe: Une nouvelle convergence?* [Christian democracies and conservatism in Europe: A new convergence?]. Brussels, Belgium: Editions de l'Université de Bruxelles.
Gehler, M., & Kaiser, W. (Eds.). (2004). *Christian democracy in Europe since 1945*. London: Routledge.
Kaiser, W. (2007). *Christian democracy and the origins of European Union*. Cambridge, UK: Cambridge University Press.
Kalyvas, S. (1996). *The rise of Christian democracy in Europe*. Ithaca, NY: Cornell University Press.
Kselman, T., & Buttigieg, J. A. (Eds.). (2003). *European Christian democracy: Historical legacies and comparative perspectives*. Notre Dame, IN: University of Notre Dame Press.
Lamberts, E. (Ed.). (1997). *Christian democracy in the European Union, 1945–1995*. Leuven, Belgium: Leuven University Press.
Mainwaring, S., & Scully, T. R. (Eds.). (2003). *Christian democracy in Latin America: Electoral competition and regime conflicts*. Stanford, CA: Stanford University Press.
Warner, C. (2000). *Confessions of an interest group: The Catholic Church and political parties in Europe*. Princeton, NJ: Princeton University Press.

CHRISTIANITY

Christianity generally refers to a system of belief that centers on the assertion that Jesus of Nazareth is Christ the Messiah, the Savior of all humanity. It is embodied in the historical reality of communities that spread across the world through 2 millennia, with a wide variety of expressions in life and liturgy. Because of this earthly embodiment of their belief in human society, Christianity has been the fount of theories and practices that are of particular interest to those in the field of politics.

Source

The "good news" that Christianity proclaims is salvation by the person and the work of Jesus Christ as attested in the scriptures. The New Testament provides basic accounts of the advent of Jesus and his subsequent reception. Because Jesus announced that his advent was the fulfillment of the promise of salvation foreshadowed in the writings of the Israelites, the Old Testament is also included in the authentic canon of Christianity. The terms *Old Testament* and *New Testament* therefore suggest a Christian perspective. In Christian tradition, the Old Testament is read in anticipation of the coming Christ and the New Testament in remembrance of the crucified and risen Christ.

The Old Testament contains the same books that Judaism views as scripture, though different in sequence and importance. It has three parts, roughly reflecting the Hebrew distinction of law, prophets, and other writings. The New Testament may be divided into two parts. The first part has four narratives of Jesus from his birth to his death and resurrection and a record of the deeds of the apostles. The Gospels according to Matthew, Mark, Luke, and John were written around the late 1st century, with different authorships, sources, and viewpoints, projecting a polyphonic but coherent figure of Jesus. The second part is composed of letters, written by the Apostle Paul and other early disciples, addressed to the scattered churches and believers of the 1st and 2nd centuries. Taken together, the Bible gives a view of the world and its history, beginning from creation to its apocalyptic consummation.

Judaism, Christianity, and Islam share not only part of their scriptural texts but also the concept of *canon*. The word is derived from the Greek word meaning *rule* and refers to a collection of writings exclusively demarcated as authoritative to their faith. The concept is missing in Eastern religious traditions, notably Buddhism, and its absence results in ever-increasing bodies of scriptures. The canon and the creeds it produces mutually authenticate each other and serve as the ultimate guarantor of "orthodoxy," a concept unique to the Abrahamic traditions. The process of editing and defining the canon took a number of politically charged stages through several centuries, and its different interpretations have often lent themselves to controversy.

Founding

Historians agree that a Jew called Jesus of Nazareth became the source of a distinct religious movement that arose from within Judaism. He was born in line with the Old Testament prophecies, proclaimed the advent of the reign of God, and was crucified by

Pontius Pilate sometime between CE 30 and 33. According to Christian doctrine, Jesus was resurrected from the dead and appeared before disbelieving disciples for a short period before his ascension into heaven. Christianity claims that this Jesus was God incarnate. Not merely was he the messenger, he was the message. Not merely did he give moral teachings and exemplify them by his life and death, he was God himself in utmost humiliation. His resurrection is understood as divine approbation and victory over death. The incarnation, crucifixion, and resurrection thus constitute the fundamental basis of Christian doctrine—that God came into this world to save humanity from their sin and alienation and restored them to wholeness by the sacrificial death of the sinless Jesus and that those who believe in him shall be saved.

The biblical accounts of this salvific event gave rise to a series of difficult questions regarding the nature of Christ and his relationship with God. The first few centuries of Christian history were spent in giving definitive forms to the faith that Christians came to believe and confess. The first Council of Nicaea, convened by the Roman Emperor Constantine the Great in CE 325, formulated the doctrine of the Trinity to confirm that Christ the Son of God is of the same substance and coeternal with God the Father. God is three persons—Father, Son, and Holy Spirit—united in one eternal essence. The Council of Chalcedon in CE 451 dealt with the issue of the nature of Christ and declared that he is truly divine and truly human at the same time. These doctrinal formulae are full of mystery not amenable to arithmetic reasoning, but they are the end result of the critical and often controversial reflections on the faith expressions of the early Christians.

The doctrinal decision-making process was naturally subject to concerns that were not purely religious, but it would be misleading to think of the development of Christian doctrines merely in terms of political manipulation and exploitation. The conspiracy theories often framed for contemporary entertainment have no place in actual history, for there is no individual or organization that has the power and authority strong enough to steer the voluntary faith of the multitude of believers in a preferred direction. Theologians and delegates of all levels, even the pope and ecumenical councils, must give due recognition to what the congregations actually believe and practice, albeit critically. The modern sanctioning of the Marian dogmas by the Roman Catholic Church is a good example. The age-old maxim *lex orandi, lex credendi* (how we pray defines the rule of faith) epitomizes this organic and indissoluble correspondence of historical faith and its theoretical formulation.

It is difficult to identify the political vision of early Christianity. While Jesus was crucified by Roman authorities for his allegedly political assertion that he was "the king of the Jews," the coming kingdom he proclaimed did not lay any immediate claim beyond individual ethics. "Render to Caesar the things that are Caesar's, and to God the things that are God's" (Matthew 22:21) well captures his spiritual claim. The teachings of both Jesus and Paul are in most cases marked with the anticipation of the impending eschaton, and as such they offer little to the constructive agenda of temporal politics. Paul's injunction not to resist but to obey political authorities (Romans 13:1–7) should also be understood within this context of imminent eschatology. Withdrawal from the world was the guiding principle for the early Christians.

History

For the sake of brevity, an overview of Christian history is presented here in three stages that combine epochs and geography. The first stage is called the "Pan-Mediterranean" era, for this is the period in which Christianity spread along the European and African shorelines of the Mediterranean Sea. The axiomatic question of this period pertains to their self-definition: "*What* is Christianity?" Using the wisdom of Greek philosophy and Roman law, Christianity defined orthodoxy by councils and dogmas and organized the church and its ruling hierarchy. After the imperial decision to employ Christianity as the state religion, it began to exercise influence in all realms of human life and extended over the European continent to form a gigantic compound of religion and culture called *corpus Christianum*. The early principle of withdrawal was replaced with that of synthesis, mutual interference, and even domination. The 11th century saw the official severance of the Eastern Church from the Western, but both traditions developed elaborate systems to balance the secular and ecclesiastical powers.

The second stage is the "Pan-Atlantic" era. A major split caused by the Protestant Reformation in the 16th century made it clear that there could be a plurality of faith in geographical proximity. For Christianity, this meant the task of designing a civil society in which religious adherence was not necessarily defined by region or clan. The paradigm shift from parish to free church was gradual on the European continent, but a small number of English-speaking Puritans took an epoch-making step by sailing the Atlantic sea to the New World to build a civil society in accordance with their own choice of religion. The Christian world evolved on both sides of the Atlantic, and their common question was "*How* should Christianity be woven into a modern civil society?" Medieval synthesis was taken apart, and new models of relationship between church and state were explored. This era coincided with the Enlightenment, and issues of human rights began to command attention through repeated experiences of persecution and oppression, providing the basis for the claims of religious toleration and other basic rights including the freedom of speech, press, assembly, and voluntary association. The modern vision of Christian politics has been reshaped accordingly, taking into account the distinction between the private and public spheres of individual life.

Christianity in the third stage moves its geographical center further West to represent the "Pan-Pacific" era. Though a thread of early missions had reached India and China in the first era, Asia came in contact with Christianity primarily via missionaries from America in the 19th and 20th centuries. In turn, it was in Asia that Christianity encountered radically different forms of religion for the first time. Facing traditions such as Hinduism and Buddhism, the predominant question now asked was "*Why* Christianity among other alternatives?" The presence of other religions relativizes the internal distinction of Catholic and Protestant, Western and Eastern Churches. As the vanguard of Christian history, Asia and the Pacific now share with the United States a leading role in the development of global Christianity. American Christianity remains exceptionally active in comparison with "post-Christian" Europe, where it is more a cultural heritage than a lived reality. In regions where Christianity shows a strong presence today, the surge of nationalism and ethnocentrism supported by conservative Christians has gained the attention of political scientists.

Contemporary Outlook

Statistics show that Christianity is no longer a religion of the West. Christians living in Europe and North America now account for a mere third of the world Christian population. Among other religions, Christianity continues to have the largest number of adherents, with Islam having the second greatest number. Within the Christian tradition, the Roman Catholic Church is by far the largest body, with more than 1 billion members. The Eastern tradition is divided by several patriarchs including Greek, Russian, and other Oriental Churches. The Episcopal Church originated in England but now has the worldwide Anglican Communion, in which the English Church is in fact a minority. Among the Protestants, Presbyterian, Methodist, Baptist, Lutheran, and other denominations have often been named "mainline," but in most parts of the world they are rapidly yielding to the fast-growing Pentecostal or Charismatic movement. This movement emphasizes direct and personal revelation through the Holy Spirit, with enthusiastic and often emotional affirmation of miraculous phenomena in efforts to imitate the fervency of the early church.

The church is a human and earthly entity instituted for the transmission and promulgation of the original revelatory content that is by nature difficult to institutionalize. As long as Christianity tries to remain faithful to its contradictory mandate, there will always be oscillation between institutionalism and the free spirit. The cycle of formation and reformation is a hallmark of life.

Anri Morimoto
International Christian University
Tokyo, Japan

See also Church–State Relationships; Religious Movements; Theocracy; Tolerance

Further Readings

Barrett, D. B., Kurian, G. T., & Johnson, T. M. (Eds.). (2001). *World Christian encyclopedia* (2nd ed.). New York: Oxford University Press.

Johnson, P. (1976). *A history of Christianity*. London: Penguin.
McGrath, A. E. (Ed.). (1993). *The Blackwell encyclopedia of modern Christian thought.* Oxford, UK: Blackwell.
Niebuhr, R. (1953). *The children of light and the children of darkness.* New York: Scribner's.
Troeltsch, E. (1992). *The social teaching of the Christian churches* (O. Wyon, Trans.; Vols. 1–2). Louisville, KY: Westminster John Knox Press. (Original work published 1912)

Church–State Relationships

Church–state relationships often produce complicated political and constitutional issues even in advanced, consolidated democracies. Despite the expectations of the modernization theory that predicted a decline in the political importance of religion with the rise of modern industrial society, current cross-national studies indicate that the relations between church and state or the question of secularism are still major issues in many societies. Even in countries where a basic consensus exists on these relationships, religiously driven moral issues, such as divorce, abortion, homosexuality, euthanasia, school prayer, and public funding of church schools and other religious institutions, are still hotly debated. This entry examines variations among Western democracies in their approach to secularism as well as the more complicated problems in the Islamic world.

The common brief definition of secularism is the separation of state affairs and religion. It is also commonly accepted that this implies the absence of an official (or state) religion (established church); freedom of religion and conscience for all citizens regardless of religion and sect; full equality before the law and access to public office, again regardless of religion and sect; the absence of a legal requirement for state acts to conform to religious rules and injunctions; and the separation of institutions performing religious services and public (governmental) activities.

Not all Western democracies meet all these conditions. In England, the king (or the queen) is the head of the Anglican Church. The Norwegian Constitution states that

> the Evangelical-Lutheran religion shall remain the official religion of the State. The inhabitants professing it are bound to bring up their children in the same. (Article 2)
>
> The King shall at all times profess the Evangelical-Lutheran religion, and uphold and protect the same. (Article 4)
>
> More than half the number of the Council of State [i.e., Council of the Ministers] shall profess the official religion of the State. (Article 12)

Under the Danish Constitution,

> the Evangelical-Lutheran Church shall be the established Church of Denmark and as such shall be supported by the State. (Article 4)
>
> The King shall be a member of the Evangelical-Lutheran Church. (Article 6)

The Greek Constitution of 1975 states that

> the prevailing religion in Greece is that of the Eastern Orthodox Church. . . . The text of the Holy Scripture shall be maintained unaltered. Official translation of the text into any other form of language, without prior sanction by the Autocephalus Church of Greece and the Great Church of Christ in Constantinople, is prohibited. (Article 3)

The Greek constitution also prohibits proselytism (Article 13 [2]) and provides for a special regime for Aghion Oros (Mount Athos), under which the region is governed by its 20 Holy Monasteries, as a self-governing part of the Greek state (Article 105).

The preamble of the Irish Constitution of 1999 declares that the Constitution was adopted

> in the name of the Most Holy Trinity, from Whom is all authority and to Whom, as our final end, all actions both of men and States must be referred . . . [and] humbly acknowledging all our obligations to our Divine Lord, Jesus Christ, who sustained our fathers through centuries of trial.

The 1997 Polish Constitution contains the following:

> all citizens of the Republic, both those who believe in God as the source of truth, justice, good, and beauty, as well as those not sharing such faith but respecting those universal values as arising from other sources.

It also refers to "the Christian heritage of the Nation" and "our responsibility before God."

The First Amendment to the U.S. Constitution states that "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof." The U.S. Supreme Court has issued many decisions that define the scope of this constitutional provision, known as the Establishment Clause.

Despite these seemingly significant deviations from the conventional meaning of secularism in the constitutional formulas, it is clear that religion no longer plays a determining role in the public life of the advanced industrial democracies. In virtually all these societies, the principle of equality before the law is firmly established, discrimination on the basis of religion is forbidden, full freedom of religion and conscience is recognized for all religions and sects as well as for nonbelievers, and the legitimacy of public (governmental) acts do not normally derive from conformity to religious rules and injunctions. This seems to be the irreducible core meaning of secularism and also an indispensable prerequisite of democratic government.

Beneath these common core elements, however, one can observe two different conceptions of, and approaches to, secularism in Western democracies. One, which may be termed *passive secularism*, essentially means a state neutral with respect to, or equidistant from, all religions and sects without precluding the public visibility of religion. "Assertive" or "combative" secularism (*laicité de combat*), on the other hand, aims at eliminating religion from the public sphere and confining it to the private domain, that is, to the places of worship and the consciences of the individuals.

Most Western democracies adhere to some version of the passive secularism model, while France and Turkey are the chief protagonists of the assertive secularism approach. The difference between the two approaches can be explained by different patterns of political development. The critical condition that affects the emergence of one or the other model, as Ahmet Kuru argues, is the absence or existence of an ancien régime that combines monarchy with a hegemonic religion. The emergence of assertive secularism in France and Turkey seems to be a response to such a combination of historical conditions. The assertive character is much more marked in Turkey because, in contrast to the strict separation of religion and state in France, the Turkish state maintains a monopoly over religious services and religious education. Therefore, Turkey provides an example not of strict separation but of a system where the individualized version of religion is at the service and under the control of the state.

The situation in the Muslim world is more complicated. It is a commonplace to argue that, as opposed to Christianity, Islam assumes the unity of state and religion and, therefore, is inherently incompatible with secularism; and to the extent that secularism is a sine qua non for democracy, this also means incompatibility with liberal democracy. Indeed, many Muslims see secularism as a creation of Christendom.

The picture is more complicated and nuanced, however, than what these extreme positions suggest. While maintaining the unity of state and religion in principle, Islam provides almost no rules in the field of public law. Even the major Islamic political institution, the Caliphate, is not based on the Koran and Sunnah (the Prophet's sayings and actions) but on the consensus of believers (*icma*) after the death of the Prophet, and it ceased to exist with its abolition by the Turkish parliament in 1924. The Islamic law (sharia) covers mostly private law relations (notably, personal status, marriage, divorce, inheritance) and provides some rules in the fields of criminal law and procedural law. Therefore, Islam has been considered compatible with very diverse forms of government in the past and at present. Many Muslim countries have secularized their legal systems to some extent, usually leaving areas such as family and inheritance laws subject to the Islamic rules. Again, Turkey is one of the very few states with a Muslim majority having a totally secularized legal system.

In addition, there remains a considerable difference between Muslim countries and Western democracies in their approaches to secularism. Even in the constitutions of most secularized Arab countries, it is stipulated that Islam is the religion of the state, the chief of the state has to be a Muslim, and the sharia is a major source of legislation. Further, certain rules of the sharia, such as those pertaining to inequality between genders, inequality between Muslims and non-Muslims, the prohibition of apostasy, and a number of rules in the field of criminal law are clearly incompatible with liberal democratic norms.

When it comes to popular attitudes toward democracy, however, there does not seem to exist a major difference between Western and Muslim societies, contradicting Samuel Huntington's (1996) clash-of-civilizations thesis. Pippa Norris and Ronald Inglehart (2004), in their worldwide survey including 11 societies with a Muslim majority, have concluded that as regards democratic values (support for democratic ideals, evaluations of how well democracy works in practice, and disapproval of strong leaders), there is minimal difference between the Muslim world and the West. According to their findings, the chief cultural difference between the West and the Muslim world concerns not political values but social values toward gender equality and sexual liberalization.

In cases of many advanced democracies, a reasonable degree of consensus over the place of religion in public life seems to have been constructed after long periods of public debate and negotiations. Even after such a consensus is crafted, however, certain religiously inspired issues are still on the political agenda. Indeed, neither the passive nor the assertive approaches are without problems. Passive secularism presupposes the existence of a broad consensus, which is difficult to maintain in countries with believers of many religions as well as nonbelievers. Assertive secularism, on the other hand, is prone to lead to conflicts between secularists and believers. Secularists will be suspicious of believers' commitment to a secular democratic system, and believers will react to secularists' efforts to eliminate religion from the public sphere.

Perhaps the key to the dilemma is the construction of "twin tolerations" (see Stepan, 2001, pp. 213–253), as the system where religious groups are allowed to organize and propagate their views in civil society and political society but refrain from imposing their views on the other members of society, respecting democratic rules and procedures, and where this boundary line is carefully guarded and monitored by the state.

Ergun Özbudun
Bilkent University
Anakara, Turkey

See also Religion; Religiosity; Religious Movements; Secularism

Further Readings

Bhargava, R. (Ed.). (1998). *Secularism and its critics*. Delhi, India: Oxford University Press.

Huntington, S. P. (1996). *The clash of civilizations and the remaking of world order*. New York: Simon & Schuster.

Kuru, A. T. (2009). *Secularism and state policies toward religion: The United States, France, and Turkey*. Cambridge, UK: Cambridge University Press.

Monsma, S. V., & Soper, J. C. (1997). *The challenge of pluralism: Church and state in five democracies*. Lanham, MD: Rowman & Littlefield.

Norris, P., & Inglehart, R. (2004). *Sacred and secular: Religion and politics worldwide*. Cambridge, UK: Cambridge University Press.

Stepan, A. (2001). The world's religious systems and democracy: Crafting the "Twin Tolerations." In *Arguing comparative politics* (pp. 213–253). Oxford, UK: Oxford University Press.

CITIZENSHIP

The concept of citizenship has a great elasticity of meaning because of its varied spatial and functional dimensions (e.g., we speak of family citizenship, state-national citizenship, cosmopolitan citizenship). The notion is all the more difficult to use as it refers to very different historical experiences, ranging from the civil makeup of ancient Rome to the contemporary experience of postnational citizenship in Europe, including numerous liberal, republican, and communitarian declensions from this membership. Another difficulty lies within an ethical framework that is an essential part of the scientific debate on citizenship. A few years ago, issues related to the ethical implications of citizenship led certain theoreticians to believe that the notion had become theoretically ineffective in social and political sciences. Although such an opinion may be somewhat excessive, it highlights the need for a rigorous definition of the concept. In this entry, citizenship is defined as (a) a juridical status of membership in a political community, (b) a status that is the condition of the political participation of citizens in the democratic functioning of the community, and (c) a status that endows individuals with a sense of citizenship that differs from their other social and cultural identities. These three basic elements of citizenship are discussed in the first part of the entry; then, recent transformations of modern civic configurations are examined.

Basic Elements of Citizenship

Three basic elements are necessary to consolidate the notion of *citizenship* and to outline distinctive characteristics that make using the term heuristically possible in political science.

Citizenship as a Juridical Status of Membership

The first element identifies the juridical foundation of citizenship as a status of membership in a political entity (often national but not exclusively so) that is juridically codified. In this regard, citizenship is often linked to nationality, which entails the recognition of rights and obligations associated with this status; in most nation-states, this criterion of nationality is still the condition for inclusion in the citizen community. With the development of migration trends and with the diversification in the scales of citizenship, the residency criterion is also sometimes taken into account to facilitate the civic integration of those who live in a nation or other juridically defined territory without full membership as a citizen of that nation-state or other group. This juridical foundation is at the center of the sociological debate put forward by Thomas Marshall when, during his famous conference in Cambridge in 1949 (see Gershon Shafir, 1998, chap. 6), he made citizenship a federative concept for political science. For Marshall, who was particularly concerned about the coherence of an English society threatened by class divisions, the interest in this concept lies in the ability of the juridical status to guarantee equal rights to all citizens. In the now established fashion, Marshall characterizes citizenship as consisting of three components: (1) the civil element, that is, individual rights such as freedom of the person, freedom of speech, freedom of thought, freedom of belief, right of ownership, right to sign contracts, right to a fair legal system, and so on; (2) the political element, including the right to participate in political activity, right to run for office, right to vote, right to petition, and so on; and (3) the social element, mainly through the free and equal access to its social welfare system. Heavily influenced by the British historical experience, this liberal and evolutionary model of citizenship still reflects an interest in putting the status of citizenship into a historical perspective that includes the birth of the constitutional state, the advent of universal suffrage, and the birth of the welfare state, all of which encouraged the progressive expansion of such a juridically codified social role. Here, citizenship is a type of indicator of political modernity. This perspective on citizenship, contested by many critics of Marshall, especially Albert Hirschmann, is based on the evolution of Western societies in which these societies moved away from feudalism and its system of discriminatory orders. Since the introduction of this argument, many works have completed this model by including new elements; in particular, scholars such as Will Kymlicka (1995) have added an important cultural dimension that establishes the notion of "multicultural citizenship," which is intended to take into account cultural, linguistic, or religious diversity. However, in *American Citizenship,* Judith Shklar (1991) shows that the symbolic possession of citizenship status is often just as important and sometimes even more fundamental than the actual practice of rights attached to it. Taking the example of American citizenship, for a long time denied to those of African ancestry and linked to the practice of slavery until the passage of the Thirteenth Amendment in 1865, Shklar argues that for many who have benefited from the fight for civil rights, what is important is not only the regular practice of rights as the recognition of a juridical and a political status that give a minimum of social dignity but also this symbolic recognition.

This first element of citizenship gives rise to a number of important debates and scientific studies in political science, including current discussions surrounding the dissociation between citizenship and nationality. Whether one considers a postnational model of civic membership, as Rainer Bauböck or Yasemin Nuhoglu Soysal suggest (see Shafir, 1998, chap. 10), or analyzes the partial denationalization caused by the emergence of European citizenship in 1992, as described by Yves Déloye (2008), the debate is centered on the pertinent degree of the social closure of the limits of citizenship and the ability today of this status to match juridically the open conception of human rights that now contrasts with the former closed-minded conception of national rights.

Citizenship as a Foundation of Democratic Practice

The second element of citizenship is characterized by the philosophical and historical link between this membership and the workings of political democracy. Because this link marks the advent of a new way of legitimizing power, which is completely dissociated from any aristocratic or theocratic perspective, citizenship entails a division of political work between the government and the governed, which has made the idea of a change in power the very principle of representative democracies. Political science literature usually compares two stances on this issue. In an elitist version, defended in the past by Gaetano Mosca, Joseph Schumpter, and more recently by Samuel Huntington, the main idea is to protect democracy from excessive interference by citizens, because civic apathy is seen as functional for the democratic political system, helping it avoid "governability crises" such as those encountered in Western democracies in the 1970s. Conversely, the participatory version advocates a strong involvement (in the sense in which Benjamin Barber uses this term) of citizens in the public sphere. The advocates of this opinion are therefore inclined to increase *acts of citizenship*. In using this term, some theoreticians have recently advocated a profound transformation of the studies dedicated to citizenship. They leave behind the perspective of citizenship as the possession of a membership status representing the modern civil freedoms guaranteed by the state, which in return demands the loyalty of its citizens, and instead put forward the practice of civic freedom and the repertoire of political expression (e.g., flash mobs, consumer awareness, hijacking political symbols through art), which play a role in making citizens the first actors of a citizenship that escapes the state's control and sometimes even questions it. The reversal of this perspective from top to bottom cannot be mentioned without referring to the old dialogue between two antagonistic civic conceptions observed by the English historian John Pocock (see Shafir, 1998, chap. 2). According to Pocock, two ideal conceptions of citizenship have coexisted since antiquity. The first, expressed by Aristotle, emphasizes the civic dimension and defines citizenship as a mode of activity specific to humanity; the second conception, outlined in the work of the Roman lawyer Gaïus, sees citizenship above all as a specific legal status and one that values its civil dimension. Beyond this founding opposition, the contemporary analysis tends to emphasize the fact that an "act of citizenship" questions the nature of "civic virtues" expected from citizens. Usually when citizenship is discussed, this questioning is related to the normative position of those who participate in the debate. The argument between liberals and communitarians is largely influential in political science and in other disciplines. Following John Rawls, liberals view citizens as individual rational agents rather than as members of a community that helps shape their values. Conversely, the communitarian critique (as presented by Michael Walzer and Alasdair MacIntyre) claims that the status of citizen presupposes an engagement in the public sphere, a sphere that constitutes the cement of the community and the solidarity that characterizes it. The topic of this intense debate, which borders on normative political philosophy and political science, points to the third element of citizenship: the impact on identity.

The Impact of Citizenship on Identity

This third element of citizenship emphasizes the principle of separation between citizen membership and other social memberships—rival sources of identification. As Jean Leca (1990) indicates in his classic text, citizenship constitutes "a civil society distinct from family, lineage and seigneurial communities" (p. 148). If the model of the Western medieval city analyzed by Max Weber is often referred to here as the paradigm of this necessary differentiation between civic identity and corporate or religious identities, it is, however, with the political experience of liberal revolutions of the 18th century that this illustration of modern citizenship develops. It is then associated with two fundamental historical movements: (1) individuation, which contributes to strongly linking individualism to citizenship, to the extent that the individual and the modern citizen merge in the *Declaration of the Rights of Man and of the Citizen* of 1789, and (2) the universalization of rights, which makes citizenship all the more inclusive. By causing a reorganization of social and religious identities between the public and private spheres, citizenship demonstrates the establishment of a social formation where the combination of status and affiliations to primary groups (family, ethnic, religious, etc.) is compatible with the promotion of an allegiance to a political community. As the classic sociologists Émile Durkheim and Georg Simmel had established, the system of division from which these civic allegiances are created favors the emancipation of the individual and the promotion of civil liberty. The classic works of Reinhard Bendix (*Nation-Building and Citizenship*, 1977) and Ernest Gellner (*Culture, Identity and Politics*, 1987) show that this civic emancipation is often in harmony with the affirmation of the nation-state, whose cultural cohesion is linked to the ability of citizenship to evoke in individuals a strong sense of identification to its culture and to its political vision. The comparative historical sociology of the nation-state, as described by Déloye (2008) and Juan Linz (1993), claims, however, that this cultural homogeneity is often questioned by the populations and therefore depends on the level of success in establishing political and social identities. One must be particularly aware of the diversity of historical paths in building the nation-state. As Linz indicates, the process of nation building—which aims to historically complete the process of state building and to develop among its citizens a strong subjective sense of belonging to the same political and cultural community—is often in the end thwarted by demands for cultural, religious, or linguistic recognition. Such demands make it politically very costly for a republican-type homogeneous citizenship to emerge, capable of making peripheral or social identities invisible—identities that individuals hold on to because of their primary socialization. Taking this historical fact into account, Linz puts forward a conceptual reflection useful for an analysis of citizenship: Establishing the exceptional character of the nation-state (with the exception of the successful model of the French state and its model of republican citizenship), he encourages us to abandon "the idea according to which every state should try to be a nation-state" in order to "turn to methods of state [and civic] integration rather than those based on national construction" (Linz, 1993, pp. 356, 365). At the heart of this conceptual shift—which leads him to prefer the notion of "state-nation" to "nation-state"—is a whole series of essential questions examined by political theory and by the comparative approach to citizenship: the recognition of cultural diversity in a liberal democracy, the development of "multicultural citizenship" models, the political construction and political validity of multinational states founded on the plurality of civic and cultural methods of identification (e.g., in Canada, India, Belgium), the "postnational" future of contemporary societies, and the transformations of the different types of political patriotism and allegiance associated with models of citizenship. The common factor here is the insistence that the observer distance himself from the abstract notion of citizenship, that of "man born again" (a favorite expression of many French revolutionaries of 1789), without the primary social or cultural characteristics, to recognize the man who is embedded, enclosed in a series of social identities that help shape his behavior in public and make him feel like a "citizen."

This shift links up with the communitarian and feminist critique formulated in contrast to a universalistic conception of citizenship linked to an "ideal" separation between the public and the private sphere. For the communitarian critique of liberalism, the issue is the atomized conception of "Oneself," with no true meaning, or historical and empirical contingency, which is capable of thinking in an independent and rational way and capable of implementing civil liberties, which are recognized by one's status as citizen in a purely autonomous way. Contrary to this definition, which Isaiah Berlin strongly defends in his *Four Essays on Liberty* (1969), the communitarian theoreticians developed the hypothesis of "Oneself," inevitably "embedded" and dependent on his common environment. This "communitarian constitution of

Oneself," to borrow Charles Taylor's terms, leads the communitarian critique of liberal citizenship to emphasize the fact that the political and civic identity of citizens is constituted on the basis of an attachment to one or several communities of reference that noticeably reduces the degree of possible differentiation from the political city.

This tension of citizenship between private man and public duty, formerly mentioned by H. Mark Roelofs (1957), is also at the heart of the feminist critique of liberal and republican citizenship. According to Carole Pateman (1989), quite rightly, this critique betrays the androcentrism of the traditional notion of citizenship, which, behind a conception of the abstract universal, tends to reflect the civil evolution of men more than it does of women. From this point of view, it is obvious that the famous typology of civic rights formulated by T. H. Marshall and discussed earlier hardly concerns women who, in most Western countries, are denied the civic rights considered by the British sociologist. In the same way that universal suffrage was considered "universal" for a long time when only men had the right to participate in the election of parliamentary elites, the notion of citizenship remained for too long a prisoner of a conception of politics that overlooked gender. In light of gender, studies on citizenship are now enriched with new sociological analysis and major conceptual reappraisals; the idea that the public and the private sphere constitute two hermetic and distinct spheres, a hypothesis that is at the foundation of liberal and republican conceptions of citizenship as described, for example, by Engin Isin and Bryan Turner (2002, chaps. 8 and 9), has been criticized because it was the main source of the exclusion of women from the political city. In a similar way, the universal perspective of citizenship was strongly modified to recognize the multiple irreducible differences between men and women. Some of the most substantial studies in gendered political science, however, deal with the inclusion of women in parliament. Whether in terms of correcting the flaws of representativeness, adjusting the electoral system to increase the presence of women in elective assemblies, or thinking of the effects of parity on politics, research on civic representation has been greatly expanded as a result of a gendered approach that most often combines political theory with comparative political analysis.

Transformations in the Civic Community

Beyond these theoretical debates, the concept of citizenship is an excellent reflection of contemporary political metamorphoses, as Isin and Turner (2002) have shown. The controversies enliven the numerous studies on citizenship that force political science to put the political actor in touch with themes such as rights and obligations, political loyalty in political configurations made more complex by the globalization and the uprooting of politics, and new methods of political communalization in multicultural societies often marked by postcolonial demands. Without trying to cover every civic transformation, Figure 1 summarizes the two main transformations of political citizenship observed at

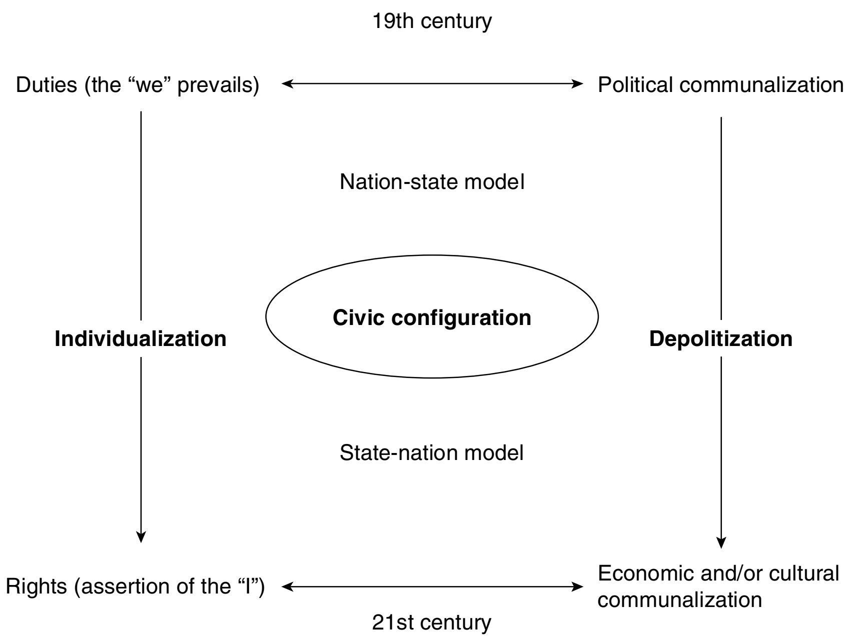


Figure 1 Historical Changes of Civic Configurations

the beginning of the 21st century. The first line of evolution concerns the normative impact of citizenship (left-hand side of Figure 1); the second transformation applies to the degree of politicization of the contemporary status of citizenship (right-hand side of Figure 1).

In the same way as classic works approach the topic of citizenship, the identification of individuals as members of a political community (local, national, or postnational) is historically linked to the normative integration of citizens. It is because individuals share not only the same system of values and representations but also common rules of behavior that they become aware of their "unity." The strength of the idea of identification with a national "we" revealed by various political science studies is also due, in one part, to this normative tonality of membership in a nation-state and its emotional capacity to unite citizens around a body of central and shared values and rules. New civic methods experimented with in Europe, North America, and Asia are probably in line with a different perspective. In many current political configurations (e.g., European Union, India, Canada, Australia, South Africa), a plurality of values coexist in the public sphere and set the social actor against cultural subsystems of different values. In this case, the question of political legitimacy is no longer asked in relation to ultimate values but often on an administrative and a technocratic level. The weakness of cultural integration—which was at the heart of this state-national civic project, for instance, the republican one—is compensated for by the economic capacity of the social welfare system to ensure to every person a certain equality of access to material well-being. The pragmatic running of society occurs here without reference to the founding values on the basis of which citizens struggle to come together because of former or present distributive conflicts. This new paradigm creates a model of citizenship, slightly inclusive in political and cultural terms but also capable of encouraging equal—more or less fair—access to economic and social goods and services (bottom of Figure 1). From this point of view, it might seem to have an affinity with historical logics used in this process of building the "state-nations" referred to by Linz (1993). However, such a conception of citizenship makes it very difficult to develop a strong sense of belonging to contemporary political spheres whose economy is also highly globalized and increasingly discriminatory.

In the manner of an ideal type, it is possible to compare these two models of civic configuration. On the one hand is the model of citizenship, which was experimented with for Western countries (top part of Figure 1), encouraging the politics of vertical integration and a certain level of homogenization or at least a certain level of cultural convergence. This model is based on a common body of values, beliefs, and representations, which conveys norms and duties that the state has to uphold. At the other end of the scale is a model of multicultural citizenship, whose horizontal inclusive logic is limited to ensuring that every citizen has equal access to the economic and social welfare of society and is free to choose his or her cultural and identity affiliations (bottom part of Figure 1). Numerous intermediate positions are possible between these two models.

A second line of evolution takes shape in the background once attention is turned to the contemporary sources of civic "communalization." Three dominant sources of communalization can be identified in a broad outline, thanks to Max Weber's old typology: a cultural source (religion, e.g., for feudal societies), a political source (e.g., the civic loyalty demanded by the nation-state), and an economic source (e.g., access to the economic market). Historically, these three sources of communalization operate in a mixed way but in varied proportions and according to different connections. What is at stake here is clarifying the respective contributions of the political order, of the cultural order, and of the economic order to the communalization of citizens. Today, in Europe and elsewhere, the market and economic or professional mobility hold a decisive place in the processes of regional integration, and these forces may lead to rivalry between groups with different cultural identities. At the same time, other social institutions, such as religion or the family, may lose some of their power to regulate relationships between cultural groups. Thus, these contemporary transformations of citizenship call into question the very status of the political with respect to its role as a source of integration and regulation. Whereas citizenship was once seen as a vehicle for distinguishing civic membership from other social identities, today it must accommodate itself to societal demands. This relative "depoliticization" appears to be the result of a new stability between different sources of civic communalization that clearly reduces the importance of political loyalties in favor of both transnational patterns of material consumption and self-assertion on a cultural level.

It is in this overall historical transformation that it seems necessary to locate present conditions of identification and of political participation in contemporary societies. It is also by taking into consideration these profound transformations that the concept of citizenship can remain a central notion in political science.

Yves Déloye
Université Paris 1 Panthéon-Sorbonne
Paris, France

See also Civic Culture; Civic Participation; Collective Security; Democracy, Theories of; Ethics; Feminism; Liberalism; Multiculturalism; Nation Building; Nationalism; Representation; Republicanism; Welfare State

Further Readings

Bulmer, M., & Rees, A. M. (Eds.). (1996). *Citizenship today: The contemporary relevance of T. H. Marshall*. London: UCL Press.

Déloye, Y. (2008). State, nation and political identity at the time of European Union. In M.-S. Darviche & W. Genieys (Eds.), *Multinational state-building: Considering and continuing the work of Juan Linz* (pp. 91–118). Montpellier, France: Pôle Sud.

Isin, E. F., & Nielsen, G. M. (Eds.). (2008). *Acts of citizenship*. London: Zed Books.

Isin, E. F., & Turner, B. S. (Eds.). (2002). *Handbook of citizenship studies*. London: Sage.

Kymlicka, W. (1995). *Multicultural citizenship: A liberal theory of minority rights*. Oxford, UK: Oxford University Press.

Leca, J. (1990). Individualism and citizenship. In P. Birnbaum & J. Leca (Eds.), *Individualism: Theories and methods* (pp. 141–189). Oxford, UK: Clarendon Press.

Linz, J. J. (1993). State building and nation building. *European Review, 4*, 355–369.

Pateman, C. (1989). *The disorder of women: Democracy, feminism and political theory*. Stanford, CA: Stanford University Press.

Shafir, G. (Ed.). (1998). *The citizenship debates: A reader*. Minneapolis: University of Minnesota Press.

Shklar, J. N. (1991). *American citizenship: The quest for integration*. Cambridge, MA: Harvard University Press.

Civic Culture

Civic culture is a subtype of political culture, that is, a set of political attitudes, habits, sentiments, and behavior, related to the functioning of a democratic regime. It implies that although citizens are not necessarily involved in politics all the time, they are aware to a certain extent of their political rights and also of the implications of the decision-making processes that affect their life and society. Both political awareness and participation are supposed to be relevant to the stability of a political regime. By contrast, citizens' withdrawal from political life has consequences not only for their ability to get what they want from the political community but also for the quality of democracy. Civic culture involves, therefore, some level of perception of the republican character of modern politics and adds a psychological dimension to the concept of citizenship. In the following, the evolution of this concept, its criticism and limitations, and its relevance for contemporary politics are discussed.

The concept of civic culture is part of a long tradition of thought that investigates the nature of democracy from a historical perspective. It refers to the role of political traditions, values, and culture in the achievement of democratization and the stabilization of a regime. Its rationale goes back to the thinking of ancient political philosophers such as Aristotle, but in modern and contemporary times also, Niccolo Machiavelli, Baron de Montesquieu, Alexis de Tocqueville, John Stuart Mill, Max Weber, and Norberto Bobbio, among others, have discussed whether a set of specific political attitudes, convictions, and behavior are a necessary and/or sufficient condition for the success of modern democracies. The question is controversial, but it has never disappeared from the debate about the necessary conditions to achieve the "good government," that is, a political regime committed to the ideal of full human realization.

While Aristotle and some Roman philosophers were mostly concerned with the excesses of the original models of democracy and popular republics, Montesquieu and Tocqueville focused on the links between the "spirit" of political institutions and laws and the "habits of the heart" that drive people toward greater public cooperation and associational life. Aristotle spoke not only of the political goodwill necessary for a political regime to be able to fulfill its mission but also of public virtues such as civic partnership and political restraint. Tocqueville claimed that the success of modern democratic regimes depended on an [individual's] self-interest rightly understood" and asserted that an open and free public life is not a natural given but depends on human will; he also emphasized the importance of a sense of moderation and self-restraint in the process of political participation.

Contemporary concern with civic culture is due mainly to two political phenomena of modern times: first, the excesses of political revolutions (e.g., the French, the Paris Commune, and the Soviet), and second, the collapse of democracy in Europe in the period between World Wars I and II, as was the case with the Weimar Republic. Democracy cannot fulfill its promises if there are no democrats; and if the nature of the democratic regime does not create an excess of political participation—which may degenerate into violence against minorities and opponents—neither does it predicate a cynical or apathetic public, which withdraws from public life and which could lead to abuses of power. As a system of government both *of* and *for* the people, but not directly *by* the people, representative democracy demands public involvement, political participation, and vertical, horizontal, and social accountability. The existence, independence, and autonomy of civil society are integral parts of it. Moreover, a widespread tolerance for a plurality of views and interests and a widely distributed sense of political efficacy, trust in government, and mutual trust among the citizenry are seen as key elements in effective civic culture.

Civic Culture as Political Culture

The most important contemporary contribution to the development of this theme appeared early in the 1960s in the work of the American political scientists Gabriel Almond and Sidney Verba. Their book *The Civic Culture* demonstrated the connection between specific political attitudes and the historical experience of democratic stability or the absence of it in five regions: Germany, Italy, Mexico, the United Kingdom (UK), and the United States.

In the context of the process of decolonization and redemocratization that emerged after World War II, Almond and Verba were concerned with the broader question of the emergence of a new world culture involving positive and negative political tendencies in countries that had become democratized as a result of external pressure, such as Germany, Italy, and Japan, and also in Third World nations recently freed from their colonial status but not necessarily from their traditional political heritage. They wanted to know whether those countries had the necessary conditions to give birth to effective democratic regimes or whether old authoritarian or antidemocratic values would prevent those countries from establishing free, competitive, and effective political systems. The continuing existence of a number of authoritarian regimes in Europe (Spain, Portugal, and Greece) and similar experiences in Latin America, Africa, and Asia in the 1950s, 1960s, and 1970s showed the relevance of their concern.

According to the logic of political culture theory—in which some critics found that psychological approaches had too great an influence—the effective performance of a democratic regime would require some kind of accommodation between the elitist conception of the role of political leaders and some level of citizens' involvement in the realm of politics. While it was later defined in terms of the strategic role of political elites in making governments govern—through their political initiatives and performance—the new theory maintained that there was also an area of collective political convictions and modes of participation to be taken into consideration. The latter was said to provide the context for the emergence of the former. The microdimension of politics, perceived at the individual level, should be congruent with its macrostructures, seen at the national level. Incongruence of political attitudes and the behavior of the masses and elites, on the one hand, and the functioning of democratic institutions, on the other, could produce a situation in which there were incomplete, untrustworthy, or dysfunctional political systems. Looking at the survival of democracy in Britain, the United States, and some European countries in the interwar period, Gabriel Almond and Sidney Verba discussed what kind of citizenry was needed to achieve a proper balance of power and to enable a stable democratic regime to be responsive.

This theory held that to work according to its principles, the democratic process should involve a healthy tension between civic obligation and actual civic performance. It required appreciation by the public of the virtues of democracy as compared with its alternatives. To submit to the law, to the mechanisms of its enforcement, and to the authorities, citizens should give both formal and informal consent to the functioning of the political system; but at the same time, they should have a proper sense of their own duties. In this respect, both the subjective and the objective political competences of citizens were considered to be of great importance. Such a political culture depended also on the levels of mutual trust among citizens. This long-established notion referred to the fact that interpersonal trust is necessary to stimulate common action directed toward the realization of collective goals. Citizens should acknowledge their obligation to be part of the system and also believe that political institutions would be accessible to their participation, but at the same time, they should be selective if their performance fell short of what was required. Then, how could one balance political activity and passivity while allowing political leaders both to exercise power and to respond to the demands and preferences of citizens? A political division of labor—strongly criticized by both communitarian and republican defenders—was seen as an essential condition to allow the political system to work effectively, to make political choices, and to encourage voters to evaluate the performance of political leaders and governments.

Politics presupposes conflict and antagonism between parties and groups in open democratic systems, but these forces should be contained by a general national loyalty and support for the political system. The civic culture approach presupposes that a political culture congruent with a stable democracy involves a high degree of consensus concerning the legitimacy of democratic institutions and the content of public policy.

Political culture has been defined as a set of attitudes, beliefs, and sentiments that give order and meaning to the political process and provide the underlying assumptions and rules that govern behavior within the political system. Expanding on this definition, Almond and Verba distinguished between three different types of political orientation: (1) *cognitive*, referring to knowledge of and belief in the political system, its roles, incumbents, inputs, and outputs; (2) *affective*, translated as feelings and sentiments about the political system, its roles, personnel, and outcomes; and (3) *evaluative*, based on judgments and opinions about political objects, particularly, the input and output processes, incumbents, and their performance.

The political culture of the five nations that they compared was characterized in terms of two key attitudinal variables: commitment and involvement. The first looks at individual attitudes toward the political system, distinguishing between *allegiant*, *apathetic*, and *alienated* political orientations; the second measures attitudes toward participation, differentiating between *parochial*, *subject*, and *participant* attitudinal types. Civic culture would combine both a participatory and a deferential perspective of politics, forming a mixed political culture in which action and nonaction, obligation and performance, and conflict and cooperation would be balanced and combined. Citizens would have to mobilize certain particular civic virtues in order to be able to evaluate issues relevant to them and their society. These civic virtues would then help prevent both the overload of the political system and an excessively deferential polity, which could give way to new forms of authoritarianism.

Critical Views of the Concept

The concept of civic culture had a considerable influence on the development of political analyses in the late 1950s and early 1960s. Conceptually linked with the so-called behavioral revolution, its proponents wanted to signal a movement away from the study of formal institutions—typical of the so-called old institutionalism—toward the study of informal behavior to see politics in real life. In the 1970s, however, the cultural approach fell out of academic fashion and came under strong criticism for being conservative, static, and tautological; for ignoring real-power relations; and for being incapable of explaining social, political, and cultural change. The view of political culture and democracy presented in *The Civic Culture* rapidly became the subject of considerable and rather heated debate. Its critics said it was deterministic, ideological, and inconsistent with the real nature of liberal capitalistic societies and only represented an idealistic (and ideological) version of politics in the United States and the United Kingdom, their prototypes of civic cultures. But the strongest charge against the theory concerned the direction of causality in the relation between culture and political structure. The British political philosopher Brian Barry and the American political scientist Dankwart Rustow claimed that Almond and Verba failed to explain the assumption about the congruence between political structure and the culture needed for democratic regimes to succeed. The culture of politics is a consequence, not a cause of institutional structure, they argued. People believe in a system through learned processes based on "habituation" and "accommodation" with the logic of democratic institutions and not the other way round. But neither Barry nor Rustow did any empirical research to prove their point. At the time he was making his criticism, Barry thought that there was not enough research to allow consistent conclusions about the controversy to be drawn and suggested that to advance the debate, scholars should study the phenomenon of transformation of political regimes to be able to grasp the real sense of causality—at the time, however, there were no regimes undergoing such a change.

Political values and orientations certainly cannot exist in a vacuum, but critics were missing the point—as Arendt Lijphart pointed out. Almond and Verba allowed causality to work both ways; in other words, political culture not only influences institutional design, but it is also influenced by institutional quality and its functioning. Moreover, the critics did not explain the normative basis of institutional choices, something else that was not to be taken as a given. In other words, institutions are chosen to carry out specific functions in order that societies can work as people want them to, and unless one assumes that institutional choices are disembodied and independent of human desires, ideals, and objectives, they do refer to normative goals, which are part of political tradition and culture. Individuals' continuing commitment to political principles—such as the right to contest power or participate in the political process—is a central aspect of the process of institutional choice. Rational motivations involved in such choices are contextualized by power relations, political culture, and tradition.

The concept has also been criticized for its supposed inability to deal with political and cultural change. Citizens are not automatons, passively receiving and internalizing political values and norms, said the critics, and although political culture is transmitted from generation to generation, it is not transmitted unchanged, nor is it transmitted randomly or without question. Cultural transmission is an active and responsive process, which is continuously being negotiated by individuals. The adult political experience over time, rather than just political socialization during childhood, is crucial to explain the ways political views are shaped. This means that experiences with institutions do matter, but critics of *The Civic Culture* never acknowledged that its authors did actually draw attention to the fact that previous political analysis had underemphasized the crucial role of experience within political systems. The study was also criticized for not properly considering the process of competing norms and values within a society. Marxist critics or liberals such as Carole Pateman also claimed that they did not deal with class cleavages in their study and were therefore not able to grasp the processes of political cultural changes occurring as a consequence of the dispute for power and political hegemony in contemporary societies.

Between the 1970s and the 1990s the civic culture approach became squeezed between two extremely prestigious and critical perspectives of contemporary political science: neo-institutionalism and the different theories that investigate the success of a democratic regime on the basis of the political and economic performance of governments, political leaders, and parties. While the first excludes any reference to a normative approach to explain the design and functioning of institutions, the second emphasizes the effects of economic and political performance. Both approaches contradict the notion that an individual's internalization of values and cultural norms is relevant in explaining the outcomes of a political system.

Consistent responses to those critics appeared, however, in the mid-1990s when the complexity of processes of political transition in different countries around the world demanded more sophisticated explanations, recognizing factors other than only those relating to institutions or political performance. The question now was the quality of democracy, not just its mere existence.

The Renaissance of Civic Culture

Recent explanations of democratization are once more returning to the cultural approach. These studies are less deterministic and are based on a probabilistic perspective about the influence of political culture. They also assume that political structures and civic culture have a mutual influence. One of its main motivations is the perception that the simple transformation of institutions that usually characterizes the end of processes of political transition is not enough to explain the different outcomes. The fact that in different parts of the world there are democratic regimes that do not fully uphold the rule of law, the principles of civil and political rights, and the mechanisms of accountability has directed researchers' attention once again to the role of attitudes, convictions, and behavior of people in explaining the variance of the quality of democracy.

This renaissance of the civic culture approach is principally due to two kinds of contributions. The first is the work of Robert Putnam on the role of social capital in social and political development. Social capital—bonding, bridging, and linking—is defined as networks and social norms of trust and reciprocity that strongly connect individuals to the common interests and goals of their community. Putnam has shown with his work on Italy and the United States that social networks based on mutual trust and the desire of individuals to act in common is decisive for the achievement of social goals and for the development of a democratic institutional environment. While in the south of Italy traditional ties do not work toward achieving openness in public life and form part of a cycle of political abuse, corruption, and governments characterized by low performance, in the north a long and cumulative tradition of cooperation and mutual trust has engendered higher levels of social and political participation and allowed regional governments to perform much better. Social trust derived from the accumulation of social capital makes citizens more cooperative, reciprocal, and willing to associate and act collectively. Thus, Putnam's contribution has emphasized how different objective and subjective conditions affect institutional development.

The second refers to the extensive work of Ronald Inglehart about the relevance of postmaterialist values from a human development perspective, which involves the consolidation of attitudes that contribute toward democratic structures. Inglehart revised the theory of modernization and argued that social and economic changes profoundly affect the traditional political culture of contemporary societies in such a way that individuals leave behind, as time passes, their attachment to survival values and develop means of self-expression that enable them to have more autonomy and independent patterns of relations with political authorities. Basing their studies on the five waves of the World Values Survey, Ronald Inglehart and Christian Welzel tested, on the one hand, the impact of values over time on the existence of effective democracies and, on the other, the influence of democratic institutions and the length of time a system of government with democratic values is in place. According to them, the results clearly demonstrate that while values are decisive in determining the existence and duration of democratic institutions, the latter have only a weak influence over time in creating a civic culture; indeed, under the influence of other variables their effects disappear altogether. They claim that studies have confirmed some basic assumptions of theories of political culture even when other variables are included in explanations of democratization processes.

Toward the end of the 1990s and at the beginning of this century, the works of Pippa Norris, Richard Rose, and Doh Shin also became associated with the resurgence of the civic-cultural approach. Using different arguments, they reacted to the analyses that in the 1970s and 1980s were deeply concerned with the decrease of political trust in Europe and the United States and also to the enormous difficulties faced by newly democratized countries in creating effective democracies. In the book *Critical Citizens* (1999), Norris and her collaborators combine the institutionalist approach with the recognition that citizens' political experience is crucial in explaining their attitudes and behavior. Shin, Rose, and others argue that ordinary people's support for a democratic regime is a learning and relearning process in which both their values and their normative perspectives have a role to play as well as their rational perception and evaluation of the functioning and performance of institutions. In both cases, without opposing the perspectives of civic culture and institutionalism, they emphasize the importance of the adult political experience of citizens in explaining the conditions under which normative expectations about the role of institutions interact with rational evaluations of institutional and governmental performance. The critical view of citizens is interpreted as part of the new civic culture.

Conclusion

The Civic Culture remains a model of scientific inquiry, allowing researchers to test, refine, or reject their hypotheses. According to Aaron Wildavsky, Almond and Verba made their evidence speak on a profound theoretical issue, and their empirical work allows others to test their assumptions. Presently, the concept is undergoing a renaissance, but this does not simply imply going back to the original approach. New developments have incorporated a more flexible interpretive and less formalistic perspective, integrating both institutional and cultural approaches that explain how citizens are involved in politics and how they influence it. Civic culture is considered to be relevant not only for the stability of any democratic regime but also for its quality.

José Álvaro Moisés
University of São Paulo
São Paulo, Brazil

See also Attitudes, Political; Civic Participation; Democratic Consolidation; Democracy, Quality; Political Culture; Political Socialization; Social Capital

Further Readings

Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton, NJ: Princeton University Press.

Barry, B. (1970). *Sociologists, economists and democracy*. London: Collier-Macmillan.

Diamond, L., & Morlino, L. (2005). *Assessing the quality of democracy*. Baltimore: Johns Hopkins University Press.

Inglehart, R., & Welzel, C. (2005). *Modernization, cultural change and democracy: The human development sequence*. Cambridge, UK: Cambridge University Press.

Lijphart, A. (1989). The structure of inference. In G. A. Almond & S. Verba (Eds.), *The civic culture revisited: An analytical study* (pp. 37–56). Newbury Park, CA: Sage.

Norris, P. (Ed.). (1999). *Critical citizens: Global support for democratic government*. Oxford, UK: Oxford University Press.

Pateman, C. (1980). The civic culture: A philosophic critique. In G. A. Almond & S. Verba (Eds.), *The civic culture revisited: An analytical study* (pp. 57–102). Boston: Little, Brown.

Putnam, R. D. (1993). *Making democracy work*. Princeton, NJ: Princeton University Press.

Rose, R., Shin, D. C., & Munro, N. (1999). Tensions between democratic ideal and reality: South Korea. In P. Norris (Ed.), *Critical citizens: Global support for democratic government* (pp. 146–168). Oxford, UK: Oxford University Press.

Rustow, D. (1970). Transitions to democracy: Toward a dynamic model. *Comparative Politics, 2*(2), 337–363.

Shin, D. C. (2006). Democratization: Perspectives from global citizenries. In R. Dalton & H.-D. Klingemann (Eds.), *Oxford handbook of political behaviour* (pp. 259–282). Oxford, UK: Oxford University Press.

Thompson, M., Ellis, R., & Wildavsky, A. (1990). *Cultural theories*. Boulder, CO: Westview Press.

CIVIC PARTICIPATION

Civic participation can be defined as

> citizens' access to and participation in information, decision making, and implementation of public policies broadly construed;
>
> taking part in democracy—individually or as a part of organized groups—through communication and public actions (including electoral campaigns and elections) where public interests prevail over private ones;
>
> the act of becoming involved in the political process and working to better the community; and

a means to guarantee the credibility of institutions, through articulation of citizens' demands and holding public officials accountable.

As the base concept of participation, the notion of civic participation is rooted in the normative assumption that the efficiency of any economic, management, and other social system as well as the legitimacy of democratic political systems depends on the involvement and participation of the public and does not simply derive from the output dispensed by the system.

The term *civic participation* concentrates on the different means by which the public can participate in the decision-making process within the political community. Democracy is considered to be the political regime that allows for the best civic participation, because democratic regimes are based on the principle of interaction and communication between political institutions (parties, parliaments, local government, etc.) and the people. Civic participation brings a great number of people with diverse sources of information and interests into the process of public decision making. Widespread and intensive public discussions enhance the critical assessment of different moral ideas, informing decision makers. As a result, civic engagement contributes to the quality of democracy.

This entry presents an analytic overview of the concept's academic interpretations. It considers the conceptualization and various empirical manifestations of civic participation and evokes several unresolved theoretical issues.

Conceptualization

Civic participation (*civic* is derived from the Latin word *civis* ["citizen"] and refers to involvement in the public sphere, as in the classical Greek *polis*) is an old concept that was introduced when civil life surpassed the arbitrary absolutistic monarchies or feudal powers, calling for individual freedom and the balance between individual liberty and social equality. Alexis de Tocqueville argued that participation in purpose-driven associations socialized people into a polity with more than merely egotistic desires, and he concluded that democracy worked through people's participation in collective activities, independent of the state.

The concept of civic participation was further elaborated in modernization theories (theories of social and economic development and democratization) and became one of the main vehicles of international (mostly sponsored by the United Nations) development policies in various regions of the world. As Gabriel Almond and Sidney Verba note, the idea is that prospects of socioeconomic development and democratization rely on a civic culture, participatory public attitudes, and social values. Through civic participation, people can influence social, cultural, and economic policies, and their participation extends beyond regular participation in elections. This means that decision makers must continuously take into account the interests, rights, and opinions of members of society. If participatory culture is weak and/or political institutions are inaccessible to people (or representation is biased toward the privileged), then the connection between people and politics is poor, compromising sustained socioeconomic development. Amartya Sen (1999) argues that "no substantial famine has ever occurred in any independent country with a democratic form of government and a relatively free press" (p. 152). The thrust of this argument is that politicians in democratic settings (with a minimum of respect for the freedom of expression) face the demands of citizens and have to accommodate them (including those of the most disadvantaged groups).

Robert Dahl presented an institutional model, linking participation within the structure of political institutions. In his account, political actors are individuals who rationally calculate if, when, and where they spend their scarce resources (e.g., money, time, know-how) in the pursuit of broader collective goals in public affairs. Individuals' public involvement is shaped by their individual taste and preference for political actions, contrasted with opportunities for other social activities (and the possibility to free ride, as the theory of collective action emphasizes). The institutional model of participation confines the issues with which each individual might be concerned and alleges that the scope and ways of people's participation are rather limited. Yet it claims that conditions of pluralistic democracy allow people to participate in public affairs roughly as equals. This approach views democracy instrumentally, as a bargaining process among different groups, where interests of each group are not subject to any further debate.

In normative democratic theory, the institutional model of participation is prominently criticized by Carole Pateman, who refined the concept of participation and focused it on a community or a social group rather than on separate individuals. Pateman elaborated the notion of participation, which, as a form of genuine civic action, is possible only in democracies, contrasted with totalitarian or authoritarian states, which force participation by propaganda or sheer repression. Civic action is defined as being undertaken collectively in pursuit of the general good, in line with the principle of political morality that benefits and burdens all citizens equally. The individual pursuit of self-interest is not its only and decisive driving force. In her generic account of civic participation, Pateman puts forward public discussions of common issues rather than rational strategies aiming at individual or narrow group goals. In a way, she makes a conceptual distinction between participation as an instrumental action (pursuing mostly protective functions) and as an interaction (with mostly educative functions) and clearly attaches more importance to the latter.

In the tradition of Aristotelian and Rousseauian political philosophy, the ideal of active deliberation lies at the basis of Hannah Arendt's conception of participatory citizenship, which in democratic polities rests on values and preferences that are discursively examined, acknowledged, and renewed. Active engagement of citizens in the public realm (discussing public issues and solving problems of their communities) provides people with the experience of public freedom and conveys a sense of their political agency (meaningfulness). In that respect, it is important to underline that Arendt cherished the idea of direct political participation and claimed that citizenship can only be practiced as a common action, collective deliberation, and the sharing of power. According to Arendt, relations of civility and solidarity among citizens can be established and maintained only through their engaged public participation. By contrast, Arendt argued that citizenship should not be conceptualized in terms of intimacy, warmth, authenticity, and communal feelings, since those inevitably led to the loss of public values such as impartiality and civic friendship and the blurring of distinctions between public and private interests.

Attempts to synthesize these two apparently incompatible models of institutional and communal participation gave birth to the concept of civic engagement, as described by Robert Putnam, Theda Skocpol, and Morris Fiorina. Here, engagement means being interested in public issues and getting and staying involved in social activities. Civic engagement refers to individuals' belonging and participating in various social groupings or associations. Civic engagement is grounded in the face-to-face interaction and communication among persons (in real empirical settings such as neighborhoods, clubs, schools, churches). As Putnam notes (2000), without "face-to-face interaction, without immediate feed-back, without being forced to examine our opinions under the light of other citizens' scrutiny, we find it easier to hawk quick fixes and to demonize anyone who disagrees" (pp. 341–342). However, civic engagement interactions usually (and realistically) do not evolve into any elevated political discussion, that is, they do not yield acute public debate about the future of the entire community or about the totality of social activities within the community. The conversations and exchanges here mostly concern everyday-life issues and community routines. Civic engagement generates social networks, habits of cooperation, interpersonal trust, and norms of reciprocity, which further contribute to the social capital, or generalized social trust (mostly measured by the trust in public institutions and in people in general). Subsequently, the stock of social capital might be applied to cooperative discussions of political issues and joint civic actions, and hence, it might ultimately contribute to better results in all domains of public life. Putnam (2000) observes that "in the language of economics, social capital lowers transaction costs and eases dilemmas of collective action" (p. 346).

From the historical-institutional perspective, Skocpol claims that associative activities are not by and of themselves sufficient to ensure good governance. Reminiscent of the protracted pernicious activities of various racist and xenophobic organizations, Skocpol argues that state institutions remain important in channeling and focusing public activism, bringing together different perceptions of the public interest as well as limiting arbitrary political power. Along with Putnam's observation of the decline of social capital in modern societies, Fiorina and Skocpol acknowledge that the rise of professional organizations and single-issue advocacy groups alienate many ordinary citizens from public life. The authors maintain that encouraging more participation by ordinary people is the best way to counteract these destructive tendencies of constricted participation in modern democracies.

Although civic engagement places broadly defined social participation before political participation, it is a type of participation that is aimed at the core of state institutions and public affairs. The civic engagement model helpfully addresses issues such as the scope and types of nongovernmental (voluntary) activities in society, the fields of nonprofit activity, the degree of public affairs politicization, and so on. Yet more questions arise when one studies the problems of defense of public interests, not readily concentrated in some geographical space and not united in some well-thought-out groups. These include environmental concerns, consumer awareness, good governance (transparency) ideals, and human rights, which are exceedingly dispersed and usually opposed by concentrated, entrenched interest groups, often including state agencies. The concept of *social movement* is frequently used to address such research puzzles.

Social movements are defined as a series of contentious performances, displays, and campaigns by which ordinary people make collective claims. Charles Tilly enumerates a series of factors that brought about modern social movements. They include urbanization (the massive presence of people in cities facilitates social interaction and the formation of purposeful groups), industrialization (significant groups of workers are concentrated in factories), mass education (higher education helps develop and articulate ideologies and collective claims), communication (technological progress makes contacts and interaction easier), and democratization (the worldwide spread of democratic regimes legitimizes the quest for political rights and empowerment).

In the contemporary world, given the uncertainties caused by representative democracy and the failures of formal mechanisms for civic participation under conditions of modern complexity, the importance of large-scale informal actions is very great. Even though informal structures and ways of participation in public affairs have always coexisted with codified procedures, yet the recent phenomenon of continuous and adaptable civic participation through social movements represents a significant expression of active civil society. New consequential societal goals (expressive and deliberative actions, devoid of purely pragmatic aims) are achieved through changes brought by social movements into existing ideological and organizational environments.

Actually, social movements extend in terms of range from local and international movements to multilevel and global movements. In terms of methods, violent movements continue to play an important role and are often shaped as semiterrorist and terrorist organizations with some social support. Yet the majority of social movements seek to achieve social and civic goals by peaceful means. Rights and obligations associated with civic participation require the constitutional framework that can be jeopardized by violence. There are important examples of (at least partially) successful peaceful social movements: Mahatma Gandhi's civil disobedience movement against the British rule in India (1930s–1940s), the American civil rights movement (late 1950s–1960s), the Polish *Solidarność*, and the Lithuanian *Sąjūdis* movements against the Soviet-imposed communist regime (1980s–1990s). Events such as the Orange Revolution in Ukraine (2004) and the Rose Revolution in Georgia (2005) can also be termed social movements because of the accepted methods of opposition and the huge popular involvement.

Concerning the interests represented in social movements, most of the 19th-century movements had their roots in social classes or cultural communities (workers, peasants, aristocrats, petty bourgeois, religious [mostly Protestant] communities, Whites, etc.). They attempted to advance their living conditions or social status. Often, these types of social movements developed into political parties. In the second part of the 20th century, new social movements (feminist, prolife, pro-abortion, gay rights, antinuclear, environmental, etc.) began to emerge and reshape the sociopolitical landscape in many countries and even internationally. The new social movements present an opportunity to reestablish the basic standards, values, and norms of civic participation. They require continuous commitment and personal activity in the public sphere.

The public sphere, in Jürgen Habermas's sense, is defined as an area of social life that has the following four primary governing principles: (1) the social status of participants is disregarded, (2) discussions within the public sphere presuppose the challenging of previously unquestioned issues, (3) issues in the public sphere are driven and decided by rational arguments, and (4) the public is inclusive. Contrary to the private sphere, where individuals decide to participate in accordance with their private interests, and to the public authority sphere, where the actions of individuals and institutions are shaped by formal decision-making procedures aimed at the implementation of public policies and the distribution of public goods, the public sphere is a sphere *suis generis* transcending individual interests and profit maximization and is mostly concerned with the deliberation of public issues.

Civic Education as an Integral Part of Civic Participation

In all these conceptualizations, civic participation is associated to some extent with civic education and the cognitive mobilization of citizens. As early as in 1861, John Stuart Mill argued that democracy tends to enhance the moral qualities of citizens because when people in some way participate in decision making, they have to listen to others, they are obliged to justify themselves to others, and they have to act sensibly, taking into account the interests and concerns of others. Rational choice theorists also agree that when people find themselves under circumstances of participatory democracy, they genuinely consider the issues of the common good and justice. Hence, civic participation tends to enhance the autonomy, rationality, and morality of participants.

In their analysis of survey data from five countries, Almond and Verba considered education (based on critical thinking and analysis) to be the most important factor creating attitudes and values vital for a participant culture. Adherents of deliberative democracy maintain that public discussions represent the only way available to democracies for the promotion of the intellectual and moral progress of the citizenry, which is their raison d'être. Proponents of the civic engagement model underline the importance of capacity development: empowering people to effectively participate in public life, to learn the routine practices of cooperation, and to acquire a taste for public affairs. Yet it must be acknowledged that people's capacities to communicate, organize, analyze, negotiate, and take positions on issues have to be commensurate with adequate capacities of staff in government institutions and other organizations in order to account for such activities.

Analytically, the extent and forms of civic participation depend on civic dispositions (interpersonal trust, people's confidence to participate in civic life, openness, tolerance, and responsibility of citizens), civic knowledge (people's familiarity with political issues and their social context; knowing one's political, civil, social, and economic rights; understanding the roles, rights, and responsibilities of citizens and public officials), and civic skills (the ability to explain, analyze, evaluate, and defend a position; using knowledge for informed participation). United Nations Development Programme (UNDP), for example, claims that sustainability, equity, and empowerment are direct results of civic education leading to inclusive participation.

Empirical Measurements of Civic Participation

Recognizing that democracy is the prerequisite for any meaningful civic participation, different forms and intensity of civic participation can be observed across countries, social groups, and times. According to Samuel Barnes and Max Kaase (1979), the level of civic participation can be seen as a continuum. At one extreme are people who have no interest in politics at all (they do not vote, petition, or demonstrate; do not read newspapers; are not informed and do not want to be informed about public issues; and never volunteer); at the other extreme are active citizens (they are interested in politics; vote; are involved in political parties, trade unions, nongovernmental organizations and civic associations; and use other forms of contention). Between these two extremes, most citizens participate through conventional electoral events (they get mobilized in special circumstances); they more or less understand public issues and occasionally attempt to have an impact on them.

Political scientists identify several sets of incentives and disincentives of participation. Pippa Norris presents an institutional account of incentives of political participation (defined narrowly as voting in elections): a proportional electoral system, not very frequent elections, small distances between the relevant contenders, and great electoral choice of candidates (parties). These factors are intrinsically related to the cognitive conditions of democracy and certainly foster civic participation. The institutional clarity of a political system and party polarization facilitate social and political mobilization, since clear ideological conflicts socialize citizens to comprehend complex political issues through simplified and normatively coherent discourses.

Putnam compiled the social capital index, consisting of measures of participation in civic and political activities, including group membership, attending public meetings on local and school affairs, service as an officer or committee member of an organization, attending club meetings, volunteer work and community projects, home entertaining and socializing with friends, social trust, electoral turnout, and the incidence of nonprofit organizations. It appears that these indicators of formal and informal civic participation and social trust are strongly intercorrelated. Research shows that local differences in the social capital index have an effect on child welfare and education, safe and productive neighborhoods, economic prosperity, health and happiness, and democratic citizenship and governance performance.

Comparative research initiated in the 1970s, such as that by Inglehart (1997), demonstrates that citizens not only want to hold their leaders accountable through elections once every 4 or 5 years but they increasingly intend to scrutinize and to influence their leaders during their term of office. It is a global evolution of continuous democracy, combining electoral rights, freedoms of association, and new kinds of contentious participation.

Another line of research considers the idea that people understand public issues and eventually involve themselves in public affairs because of and through the images given by political leaders and the media, not through personal experience. Therefore, along with institutional account and civic education considerations, research on political communication is an integral part of the study of civic participation. Civic journalism dwells on the principle that adequate coverage of politics is important for public life. The role of mass media is very important for civic participation: When journalists, academics, human rights activists, intellectuals, and artists publicly contest various norms and decisions, they safeguard the rights and freedoms of the entire community, benefiting from mobilization and vigilance of its most active citizens.

Theoretical Tensions Regarding Civic Participation

There is an inherent tension between the universalistic (communal, civic engagement, social movement) and functional (institutional) concepts of civic participation. The universalistic rhetoric claims that any public action is derivative from the civil society and contributes to the general interest, common good, and justice and not just to the agenda of interest groups. The idea of civic participation assumes that the actions undertaken by organized groups (mobilized minorities) can contribute to the benefit of the entire community and strengthen the administrative accountability and political legitimacy of the regime.

Interest groups frequently use words such as *citizens*, *civil society*, *people*, and *general public* (practicing pluralist, lobbying, or corporatist strategies). Yet most concrete proposals concern organized groups. Therefore, in empirical terms, the logic of negotiation between the institutions and some limited groups remains at the core of active citizens' participation.

Another problematic issue of civic participation has to do with its limitation to the nondecision (the consultative, predecision) stage or the deliberative part, although contemporary democratic governments claim to promote civic participation throughout the whole policy chain. Civic participation in the implementation and evaluation of public policies is limited. Considering that actual public policy decisions usually are taken by public authorities in cooperation with selected organized groups, other social actors have to face the decisions *ex post.*

The thrust for promotion of civic participation is derived from the liberal conception of democracy. Openness and transparency, organized consultation, better communication, decentralization, and so on can make institutions more accountable. However, these techniques are not substitutes for the representative forms of democracy and citizenship, and they do not encourage an average citizen to become active, since they require considerable intellectual and financial resources.

Finally, the concept of civic participation concentrates on sets of functional groups with particular ends (trade unions, employers' organizations, professional associations, community-based organizations, charities and religious groups, etc.). Yet civic participation research often overlooks the general actors (e.g., political parties) who represent a global view at the grassroots, while only the top of their organizations belong to the public authorities, acting on behalf of the state.

Irmina Matonyte
Institute for Social Research
Vilnius, Lithuania

See also Civic Culture; Civil Society; Interest Groups; Participation; Social Capital; Social Movements

Further Readings

Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton, NJ: Princeton University Press.

Arendt, H. (1977). Public rights and private interests. In M. Mooney & F. Stuber (Eds.), *Small comforts for hard times: Humanists on public policy* (pp. 95–119). New York: Columbia University Press.

Barnes, H. S., & Kaase, M. (1979). *Political action: Mass participation in five Western democracies*. London: Sage.

Dahl, R. A. (1956). *A preface to democratic theory*. Chicago: University of Chicago Press.

Habermas, J. (1989). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge, UK: Polity.

Inglehart, R. (1997). *Modernization and post-modernization: Cultural, economic and political change in 43 societies*. Princeton, NJ: Princeton University Press.

Mill, J. S. (1982). *Considerations on representative government (1861)*. Oxford, UK: Oxford University Press.

Norris, P. (2002). *Democratic phoenix: Reinventing political activism*. New York: Cambridge University Press.

Passerin d'Entreves, M. (2006). *Hannah Arendt*. Retrieved February 2, 2009, from http://plato.stanford.edu/entries/arendt

Pateman, C. (1970). *Participation and democratic theory*. Cambridge, UK: Cambridge University Press.

Putnam, R. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

Sen, A. (1999). *Development as freedom*. Oxford, UK: Oxford University Press.

Skocpol, T., & Fiorina, M. P. (Eds.). (1999). *Civic engagement in American democracy*. Washington, DC: Brookings Institute Press/Russell Sage Foundation.

Tilly, C. (2004). *Social movements, 1768–2004*. Boulder, CO: Paradigm.

Tocqueville, A. de. (2000). *Democracy in America* (H. C. Mansfield & D. Winthrop, Trans. & Eds.). Chicago: University of Chicago Press.

United Nations Development Programme, Bureau for Development Policy, & Democratic Governance Group. (2004, April). *Civic education: Practical guidance note*. Retrieved February 2, 2009, from http://www.undp.org/oslocentre/docs04/Civic%20education.pdf

Civil Service

A civil service is a system for managing human resources using mechanisms with which governments select, employ, and promote the personnel who act as public administrators. A civil service system takes into account the qualifications and aptitudes of employees. In other words, it is the institutional framework within which the human resources employed in the public sector are organized in a way that allows governments to employ the best people. These employees work to achieve the results of projects put forward by their governments, regardless of their political affiliation or party. The bases of the civil service are established in statutes, laws, or regulations, and for its successful operation, an autonomous structure is required—a body that will oversee its strict operation and the observation of the norms that regulate the civil service, so that its usefulness contributes to strengthening the democratic values of the societies it serves. This entry describes the characteristics of career civil service, the forms it may take, and the subsystems involved in its operation. The place of the civil service in public administration and the benefits associated with it are then examined in light of evolving democratic values.

Characteristics of the Civil Service

Given its autonomous role as described above, the civil service is not simply a series of rules and technical conditions; rather, it should be seen as a type of moral code for public service that is designed to achieve common objectives and should be shared by all its functionaries. Civil services are based on three fundamental values for a democratic public administration, taken principally from the logic of bureaucratic organization proposed by Max Weber: (1) equality, (2) merit, and (3) impartiality. Thus, modern civil services are based on the following:

1. Recruitment is based on the *merits* of the civil servants and not on patronage or inherited posts.
2. The idea of *equality* allows anyone to participate in the competition to occupy public posts, provided that they meet the prerequisites established for the position in question; and, once they are in the civil service, the system of evaluation and any sanctions should be comparable for all the functionaries.
3. *Impartiality* guarantees that the members of the civil service should maintain a neutral position toward political parties, related to an internal logic, to grant all applicants the same opportunities and, when dealing with the public, provide the same attention and quality in the services offered to all citizens regardless of their gender, religion, or partisan preference.

The concept of a career civil service can be given at least two meanings. First, it is used as a synonym of public bureaucracy, although not all the government functionaries necessarily form part of it. The Latin American Center of Administration for Development (CLAD) maintains in the Ibero-American Charter of Public Affairs that, in Ibero-America, the terms *public function* and *civil service* are generally used as synonyms. Within this acceptance, there is an enormous variation in the use of the term, since sometimes *civil service* may in fact consist of any or all functionaries who carry out administrative work within a government. The second use given to the term *civil service* refers specifically to the area or areas of the government in charge of performing the recruitment and evaluation of the functionaries.

Therefore, not all civil services are the same; the systems of recruitment, sanctions, and recognition and, above all, the attributes and extent of the civil service vary according to the political and institutional context they are in. It is important to mention that within the same civil service there can be various types of labor relations; this means that there can be some functionaries appointed by direct political nomination and others by means of a public competition.

Members of the civil service are known as career public servants, and they can be employed at any level of government (federal, provincial, or local), depending on the size of the civil service, or in the autonomous institutions, such as electoral organisms or any of the three powers of government: executive, legislative, or judicial. In some countries, such as the United Kingdom (UK), career civil servants are only those who work for the Crown; functionaries who serve in the lower levels of government are not included. In the United States, on the other hand, there are civil servants within state and local governments, and in France, officials of some public enterprises are also included.

Further, career civil services are regulated by laws or statutes that have different characteristics relative to the rest of the labor legislation for the country in question; as a result, these functionaries usually remain separate from other public employees. These differences become apparent in the policies relating to retirement and the establishment of pensions, as well as the type and degree of sanctions, or in the capacity to organize themselves in unions and other associations.

As the name suggests, the military are not included in career civil services, although some civil servants who work in the areas of national defense are generally considered to be career civil servants. There are also functionaries who hold their posts because they were political nominees or because they were elected by popular vote, which places them outside the civil service. In the case of Ireland, for example, government ministers and some of their political advisors are excluded from the civil service.

Types of Civil Services

Closed Versus Open Services

Conceptually, a career civil service can be either open or closed. A closed service takes for granted that persons entering the public service plan to develop their whole career there. It is a closed system because one can enter only at the lowest levels of the organizational pyramid and from there gradually climb up to the higher positions. That is, personnel tend to be hired for the lower echelons, and from there, mobility through the ranks may be possible. To enter, one has to have a certain level of education and submit to a public competition in order to be selected; because entrance takes place at the lower levels of the organization, the participants' experience has little relevance in the selection process. One great advantage of this system is the creation of an ethos in the public service, as the functionaries manage to accumulate important knowledge about their organizations. The following are the principal disadvantages of this model. The first is that upward mobility within the organization is usually linked to the seniority of a functionary and not to his or her performance. The second is that this system limits the universe of contestants to those who have accumulated little or no experience; finally, this model can produce unwanted results by generating incorrect incentives, and it can even lead to poor performance by the functionaries, by not generating competitiveness with the outside labor force. The oldest civil services are built on the basis of this model and have found it necessary to make important revisions to their schemes, as in the cases of the career civil services in France and Germany.

An open civil service, in contrast, allows for public competition not only for posts at the lower levels but also for the higher positions within the organization. The system is therefore open to outside participants, although there may be a clause stipulating that preference be given to those who already belong to the civil service. In this system, the creation of a spirit of public service becomes less evident, since there is the possibility that employees from the private sector can become public functionaries and vice versa. Nevertheless, one evident advantage of this assumption is that by alternating between private enterprise and the public sector the number of participants increases, allowing for a broader universe for recruitment of the most qualified as well as the benefit of the transfer of knowledge from one sector to the other. These employment civil services are being tried out in countries that have recently created or reformed their civil services, such as New Zealand and the UK.

Executive Versus Senior Service

Because of their scope at the different levels of government, the civil services are also classified as executive or senior civil services. This distinction implies the separation between the authority of the higher and middle levels of command within the public sector.

Operational Subsystems

Whatever model is adopted, the civil service operates with certain subsystems that are indispensable for their proper functioning. The subsystems that make up a "traditional" civil service are the following.

Recruitment and Selection

The subsystem for recruitment and selection establishes the selection procedures in the contest for the position (if there are to be quotas for positive discrimination or not, i.e., if it is mandatory to include applicants or representatives from minority or marginal sectors of society). In addition, it specifies whether the contest is to be public or closed, what type of exams and study guides are to be used, and what interviews are to be held and with which of the functionaries, among others. It also establishes the requisite profile and academic qualifications, experience, samples of work, and personal references that the candidates for the post should have, as well as establishing the conditions under which they would be integrated into the public sector, that is, what type of contract they would subscribe to, what salary they would receive, and if they would be evaluated before being confirmed in the post or if they would first be subjected to a trial period. For this procedure, first, there should be a process of revision of the applicants' curricula vitae; second, they should undergo examinations and psychometric tests; and third, they should submit to an interview, and on occasion, as previously mentioned, they are given a trial period in the post. Ideally, this subsystem should have mechanisms allowing rejected candidates to appeal the decisions made by the recruitment team, something that is not always taken into consideration. Moreover, the system demands constant revision to remain adapted to the requirements of the different posts, revising among other things the pertinence of the evaluations to make sure they correspond to the catalog of existing posts.

Education and Training

The subsystem of education and training regulates the training of the public servants, whether initial, continuous, or a combination of both. Initial education consists of generating mechanisms that allow for the development of competence relevant to the post aspired to, which would not have been acquired during the professional education of the functionary, as well as those that lead to the homogenization of the knowledge of the functionaries to facilitate more or less uniform performance in the job. It should be mentioned that not all civil services have subsystems for initial formation, as many demand that this should be provided by a particular institution, which could be internal, such as the Civil Service College of the UK, or by an external institution such as the Hautes Écoles or the École Nationale d'Administration in the case of France. Continuous education, on the other hand, responds in some cases to the demands of the public and in others to the modifications or updating of the forms and procedures of the different administrations or the incorporation of new values in the face of changing scenarios, including in organizational culture. Recently, this subsystem has greatly benefited from the advances in information technology, from the possibility of attending courses and presenting evaluations at a distance. All this has, no doubt, resulted in better use being made of working hours and improvement in the activities undertaken by public functionaries. Although one should mention that in many cases these resources are not optimized because continuous education has been focused on nominative subjects instead of operative affairs, which detracts from the possibilities of immediate usefulness.

Evaluation

The subsystem of evaluation seeks to identify the strengths and weaknesses in the public management from individual and organizational evaluations. It is necessary to decide in what domain to conduct an evaluation, how evaluation is to be carried out, what indicators will be used and with what frequency, and what is to be corrected or improved. This subsystem can be linked to sanctions applied to public functionaries, ranging from verbal reprimands to permanent removal from the post. It also covers the schemes for evaluating performance, where there is a relation between wages or the bonuses of the functionaries in relation to the goals set, in the first case, and to an outstanding performance, in the second. Linking the assessment of performance to the budget may be advantageous because it leads to increases in efficiency and effectiveness, better planning, and a clearer focus on results. However, it can also have its disadvantages, by exceeding the budgets of the organizations or by cutting resources in areas where they are needed, without determining accurately the causes of a deficient performance, which may have been, for example, due to inadequate resources. In addition, evaluations should become an important input for the subsystem of continuous education, since they can identify the areas that require a greater effort in formation and training and the contents that these options should include.

Promotions and Mobility

The subsystem of promotions and mobility manages the horizontal and vertical mobility of the civil servants. Because mobility can be horizontal as well as vertical (particularly in the case of an open employment system), the exchange and transfer of knowledge from one area to another are encouraged, although this situation can cause a loss of experience within organizations. Mobility can occur both in an open system, where anyone can have access to a post, and in the so-called elitist systems, in which only graduates from certain universities or careers can occupy positions at the highest levels. In the more traditional systems, promotion of functionaries has been related to seniority in the post and within the schemes of the civil service; however, as a result of reforms arising from the adoption of ideas associated with the new public management approach, priority is now sometimes given to the performance of functionaries rather than to the length of time that they have held the post.

Removal From Office

Under the traditional principles of administration of the subsystem for removal from office, a functionary would leave office only to receive a pension. However, the stability that this immobility gave the functionaries is a negative performance incentive. Current tendencies indicate that removal from office can occur for at least two reasons: (1) the civil servant's retirement or (2) the temporary removal or separation from office due to the poor performance of the public servant, a situation that should be very accurately documented.

Public Administrations and Career Civil Service

The successful implementation of a career civil service is indispensable for improving and modernizing public administrations, since this allows for greater continuity in public policies and also increases the efficiency and efficacy of the administration. In addition, to set in motion a professional civil service requires prior institutional action to integrate a catalog of posts that allows complete restructuring of public administration from within and in the medium and long terms. Detailed job descriptions of each post also favors an increase in accountability within each organization.

The career civil services are remarkably effective in reducing corruption and increasing the equity of the public services, since functionaries know that permanence in their posts depends on their performance and not on electoral cycles, political movements, or arbitrary dismissals, all of which discourage them from creating and attending to specific clienteles. The autonomy that the administration acquires facilitates continuity in the application of programs and laws, because the administration is evaluated by the degree of fulfillment of the programmed objectives, and it also minimizes the influence of the politicians on the implementation of programs.

Similarly, adequate administration of human resources inside government results in enhanced transparency, as much for individuals as for organizations, since beginning with the integration of a catalog of posts and a clear definition of organizational structures, one can know who is responsible for the results of programs and government policies. At first glance, civil services may appear rigid; however, when their schemes are well integrated, they can even give greater flexibility to the public administration since it is easier to make quick adaptations to the demands of changing circumstances.

The professionalization of the public administration can also result in greater confidence by citizens. First, by legitimating the government through more efficient administration, and second, by being more flexible, and thus more successful, in the interaction with average citizens. In fact, a civil service system also tends to foster values within public administration such as impartiality, awareness of financial limitations, as well as legality and integrity of the government and public officers. These values contribute to the collective construction of an ethos for the public sector, indispensable for the identification of the personnel as part of an organization that aims at the collective, consecutive pursuit of objectives, preserving the institutional memory and promoting its prestige outside the organization. The public sector ethos also has a function within the public administration: It enhances communication between the various members of the organization and can direct efforts toward common objectives. In this way, communication and coordination between bureaucratic organizations can improve. Moreover, career professional services within the civil service can build up skills and expertise within the public administration, which, in the long run, will enhance not only the government institutions but also a solid democratic system.

From their perspective, public functionaries are also directly benefited by forming a part of civil service, since apart from the obvious advantage of not being arbitrarily removed from their offices, they are subjected to clear regulations and specific responsibilities, and they receive a continuous education that allows them to keep their knowledge up-to-date and to eventually make a career within the public administration, which is useful socially.

The nature and usefulness of the schemes for a career civil service have been the subject of intense debate and important processes of adjustment because, among other reasons, they are recognized as a fundamental element in the processes of renovation and modernization of public administrations. A career civil service plays a central role in the transformations favored by democratic schemes, such as commitments to making public administration more transparent and improving the mechanisms for public accountability and the management of public resources. Professionalizing the public function is an indispensable condition for national development.

María del Carmen Pardo
El Colegio de México
Mexico City, Mexico

See also Administration; Administration Theory; Autonomy, Administrative; Development Administration; Public Employment

Further Readings

Aberbach, J. D., Putnam, R. D., & Rockman, B. A. (1981). *Bureaucrats and politicians in Western democracies.* Cambridge, MA: Harvard University Press.

Butcher, T. (2003). Modernizing civil services: An era of reform. In T. Butcher & A. Massey (Eds.), *Modernizing civil services* (pp. 1–15). Cheltenham, UK: Edward Elgar.

Geddes, B. (1991). The politician's dilemma: A game theoretic model of reform in Latin American democracies. *American Political Science Review, 85,* 371–392.

Heady, F. (1996). Configurations of civil service systems. In H. A. G. M. Bekke, J. L. Perry, & T. A. J. Toonen (Eds.), *Civil service systems in comparative perspective* (pp. 207–226). Bloomington: Indiana University Press.

Hood, C. (1991). A public management for all seasons? *Public Administration, 69,* 3–19.

Horn, M. J. (1995). *The political economy of public administration.* Cambridge, UK: Cambridge University Press.

Knill, C. (1999). Explaining cross-national variance in administrative reform: Autonomous versus instrumental bureaucracies. *Journal of Public Policy, 19,* 113–139.

McGregor, E. B., Jr., & Solano, P. (1996). Data requirements and availability. In H. A. G. M. Bekke, J. L. Perry, & T. A. J. Toonen (Eds.), *Civil service systems in comparative perspective* (pp. 42–66). Bloomington: Indiana University Press.

Peters, B. G., & Pierre, J. (1998). Governance without government? Rethinking public administration. *Journal of Public Administration Research and Theory, 8,* 223–243.

Thompson, J. R. (2000). Reinvention as reform: Assessing the national performance review. *Public Administration Review, 60,* 508–521.

Civil Society

From the perspective of political sociology, the concept of civil society refers to a certain type of society. Its basic characteristic is the attitudinal capability of its members, that is, citizens, to actively organize themselves to pursue certain (common) goals, within a framework of specific formal institutions. In detailed definitions, depending on particular research goals, civil society may therefore be defined in either attitudinal or institutional terms. This entry discusses their perspectives, their origins, and their applications to the contemporary world.

In the attitudinal perspective, the pivotal concept is that of citizenship as an individual's attitude toward the entire political collectivity and other individuals, as well as to institutions and procedures. The basis of this attitude is seen as the individual's internalization of the rights and obligations associated with membership in the collectivity, and this attitude constitutes the foundation of the citizen's role.

Elements of citizenship thus defined include, for instance, political identity and self-awareness, the character and sources of agency, law and order, attitudes toward recognized authorities, attitudes toward pluralism and diversity, the level of openness toward the terms of membership in a political community, and, finally, convictions regarding the nature of civil rights (their universality or particularity). From this perspective, a civil society is a state of democratic political culture, a conglomerate of norms and values typical of a community of citizens who are free and equal before the law, which they shape and develop through their active participation in a public life.

In the institutional approach, civil society is seen as an arena of citizens' activity embedded in an institutional framework, a structurally described sphere of social and individual autonomy that extends between the state and private life. This institutional framework includes, among others (constitutionally guaranteed), civil rights and freedoms (in particular, freedom of speech, assembly, and association and the right to vote), the rules for obtaining citizenship, the rule of law, and welfare state institutions. In this context, a civil society usually stands for the totality of institutions, organizations, and associations operating in the public sphere with relative autonomy from the state, established at the grassroots level and usually characterized by voluntary member participation.

No matter how it is defined, the concept of civil society is inextricably intertwined with the concepts of collective and individual autonomy and political sovereignty. To that extent, it also contains, at its deepest level, a strong normative postulate on how an autonomous community of free and equal citizens should be structured. This normative aspect is manifested in various theoretical approaches that present diverse solutions regarding the relations between the collective and the individual and between the private and the public. The opposing positions on that matter are especially clearly displayed in the disagreement between classical republican (or today, communitarian) concepts and liberal theories, which can in fact be described as a normative controversy over civil society. Theories of civil society, articulating the ideal of a free and self-governed community, inevitably refer to some kind of democratic social order. In the Tocquevillean perspective, civil society is a space for associational life and a condition for preserving a democratic political culture and truly free democratic institutions. Hence, this concept is cognitively useful especially in analyzing both the theory and the practice of democracy. It also gains a special importance in the era of globalization, in the modern world of cross-border structures, in the slogan of a "global civil society."

Civil Society: Classical Roots and Modern Developments

The idea of a civil society is deeply rooted in European culture. It goes back to ancient Greek and Roman thought. To Aristotle, a civil society (*koinonia politike*) was an ethical and political community of free citizens pursuing and achieving full humanity through active participation in the life of the *polis*—its tradition, law, morality, and interest. The ancient Greek *agora* as a space of communication for citizens, a forum for exchange of arguments, for sharing visions of the common good and taking communal decisions, was the proper space of cogovernance and, hence, the realm of politics. What the Roman tradition contributed to the idea of civil society was the concept of natural law, discovered through rational thought. According to Cicero, the civil society (*societas civilis*) was the highest kind of social order, reflecting the eternal and unchangeable natural law, a community of people who recognize the same law and work together for the common good (*res publica*). A continuation of such republican thought on the eve of modern times can be found in the writings of the Renaissance Italian thinkers. For Niccolò Machiavelli, the civil society was a political community bound by efficient governance that ensured its stability, internal order, harmony, security, and glory. At the same time, this was a space for political virtue to emerge and be pursued. Despite differences, these classical republican perspectives have in common the perception of civic community as valuable in itself, defining a good to which individuals' particular interests must be subordinated. Here, the freedom and sovereignty of a political community as the common good is the prerequisite for an individual citizen's freedom, which is only achievable through membership in this political community. What regulates those freedoms is the authority of the law and political institutions as well as shared values and customs. This common good is discovered, identified, and, as a consequence, pursued through citizens' participation in public life. In this classical republican perspective, civil society is therefore a moral and political community pursuing a common good, where its citizens realize their freedom by exercising their rights, fulfilling their duties, and exhibiting civic virtues.

Modern times brought a dramatic reversal of this relationship. The revolution started with rationalism as a philosophical trend that regarded each human being as capable of rational judgment and established the individual, rather than the community, as the source of morality. This thought was followed up by the pioneers of the Enlightenment and liberalism, who held equal rights of individuals (e.g., habeas corpus, private property) to be natural and inalienable. The primordiality of those rights vis-à-vis the political assembly is expressed in the concept of the "social contract," according to which a voluntary agreement of free and equal individuals is established to secure their natural rights against the abuse of power. The despotism of nation-states that emerged after the Westphalian Treaty of 1648 often gave rise to a negative definition of a citizen's freedom as a sense of security (founded on a law that restricts the arbitrariness of other political actors). Subsequent political revolutions in England, the United States, and France adopted and strengthened this modern perspective of an individual as holding natural rights and freedoms that a state, as a legal institution established by a voluntary political assembly of its citizens, must safeguard.

The classic liberal thinkers believed in fundamental individual rights that include property rights and the freedom to dispose of one's property. According to liberalism, the basic domain of freedom is the market, through which individuals can most efficiently pursue their particular interests. The clash of those individual interests is seen as a universal model shaping the dynamics of the free social space. Hence, in this liberal perspective, civil society is treated as a universe of free individuals who pursue their own goals in their own way while respecting agreed-on principles. It is a sphere of independence and freedom extending between the family and the state, where a spontaneous self-organization of individuals emerges, making it pluralist and heterogeneous (politically, economically, and culturally). What immanently binds this social space of diversity and conflicting interests is, according to the followers of the Scottish Enlightenment, either the inherent social impulse—the friendship and altruism of human beings—or conformism, vanity, and a desire to please others. Even though distinguished from the institution of the state, the sphere of civil society remains in a constructive relationship to it, as citizens use it to decide with others on matters of shared interest.

This liberal perspective on civil society became a negative point of reference for thinkers such as Georg Hegel and Karl Marx, who did not share its optimism as to the immanent powers regulating the conflicts of particular interests and egotisms. Both Hegel and Marx explicitly equated the civil society (*bürgerliche Gesellschaft*) with the economic society and the private sphere, with which they contrasted the public sphere as the only one capable of curbing conflicts and disparities within civil society. According to Hegel, the embodiment and guarantor of this superindividual order was the state. For Marx, it was a political civil society that can be created after the abolition of private property, which is the source of disparities, domination, and social conflicts. It was with reference to these two thinkers that Antonio Gramsci developed his concept of civil society in the first half of the 20th century.

However, from the mid-19th century until the 1970s, the interest in the concept of civil society virtually disappeared, replaced by a more detailed reflection on the institutional and normative framework of the democratic order. The renaissance of this idea in the last decades of the 20th century is related to the birth of opposition movements in Latin America and Central and Eastern Europe, where self-organizing social and civic structures operated beyond the state or in opposition to it. From the perspective of mainly Western analysts, the category of a civil society seemed to be an appropriate tool to describe them. Civic movements in authoritarian countries were undoubtedly a major impulse for the renaissance of interest in the category of civil society. It was all the more important, however, that they occurred concurrently with social and political phenomena and changes taking place in Western societies themselves (e.g., the welfare state crisis, new social movements). In the public debate, criticism was directed either at the excessive welfare state or at the alienation of representative democracies from civil society. The advocates of social and economic mobilization or social and political democratization, of increasing civic participation and involvement in public life, and of combating alienation and individualism quite often made reference to the classical republican visions of the truly civil society. At the same time, they forced political thinkers to revisit this issue in the context of the modern era.

With respect to the concept of civil society, the most interesting aspect of this debate is the dispute between liberals and communitarians, pursued with varying intensity over the past 30 years of the 20th century. Even though it is rooted in two different visions of the individual–collectivity relationship and the factors constituting the Self, this dispute finds its practical expression mostly in the different understanding of a civil society and its relationship to the democratic state. Putting it as simply as possible, for liberals, a state is one of the many voluntary assemblies of free and equal citizens who agree on the principles of justice, which they can reflect on, question, and renegotiate. Its basic role is to safeguard the agreed procedures. Those procedures (e.g., the rule of law) and the results of their application are what liberals are inclined to consider a "procedural" common good. A state, largely restricted to safeguarding the agreed rules, would open even more space for the development of a civil society. As a guardian of procedures, it should remain neutral to various goals, visions, and ideals of good life, which are to emerge and compete with one another within the very space of the civil society, in the "free marketplace of ideas." It is civil society rather than the state that selects and orders these procedures. In this sense, a liberal civil society, as distinct from the state, remains in a constructive relationship with the state but is more significant than the latter. A political society is important but has no absolute priority.

For communitarians, by contrast, the state is the most powerful political community, defined as a particular "community of experience" in which a shared past and future are intertwined. The community's tradition is its substantial common good, and its maintenance is a common goal and endeavor. Active participation and involvement in the community's life leads to identification with its goals and generates a sense of solidarity, so promoting the culture of participation is a value worth supporting. A state may remain neutral on various matters but not on the matter of patriotism. Only by discovering the principles, ties, and loyalties that bind the community together can individuals fully understand their social and civic roles. Civil society is not a chaotic set of competing or cooperating associations but an entity united by a shared self-awareness of its own history, tradition, culture, institutions, and achievements; the goals and concepts presented in the "free market of ideas" must be intelligible to everyone and must be embedded in the culture or tradition of the community. Thus, the state is entitled and obliged to support a common understanding, by following the policy of preference for the common good and establishing a hierarchy of goals and values to inspire its members. This policy of preferring a substantially understood common good over the policy of neutrality makes the state the ultimate entity. In this sense, the state is more important than the civil society, or to put it in other words, a civil society is tantamount to both the political community and the state.

This controversy over ideas, theories, or even ideologies can often be found in daily life and disputes in liberal democracies. It refers, among other things, to the practical dimension of the relationship between the state and the civil society, in particular in different models of cooperation between public administrative structures and the so-called third sector.

State and Civil Society

Analytically, we can distinguish three types of relationships between the civil society and a democratic state. In the first, a civil society is a complement to the state. In the second, the state is a complement to civil society. In the third, the state and civil society are in an antagonistic relationship. The first type of relationship emerges where the state is the major agent for the redistribution of national wealth and takes responsibility for providing and operating many spheres of collective life (e.g., education, health care, pensions). In this situation, civic organizations enter only those areas of collective life that the state does not provide with sufficient social services. Such a relationship can be observed in welfare states characterized by an extended sphere of social benefits, and its theoretical grounds can be found in communitarian concepts. An excessive welfare state may, however, lead to the "learned helplessness" syndrome (or dependency on social benefits), which converts some citizens into the welfare state's clients. This weakens the vigor of civil society.

The second type of relationship refers to the situation where the state's responsibilities are restricted to the minimum (e.g., police, army, courts, diplomacy). The remaining part of public life is taken care of by grassroots citizens' groups. Such a relationship stems from liberal concepts of citizenship and state. It is worth noting that where a state's noninvolvement is too restricted, social inequalities may increase, which may lead to the marginalization or exclusion of some segments of society. This, in turn, reduces the vitality of a civil society because for marginalized individuals, dealing with public matters ceases to be of importance, having to focus on survival. The third type of relationship echoes the relation between individuals and the authoritarian state. It can be found predominantly in young democracies as this antagonistic relationship between citizens and the nondemocratic state is still present in the collective memory and in common cognitive structures. In this case, various organizations of civil society usually have little trust in the state institutions (and vice versa). As a consequence, conflict prevails over cooperation. Over time, especially when a young democracy has completed the consolidation phase, this generalized antagonism may disappear, replaced by civic defiance focused on certain acts of a democratic state.

Cognitive Usefulness of the Concept

The usefulness of the concept of civil society changes over time. There are periods where researchers' attention is drawn to other areas, with the civil society itself being treated as secondary to some more fundamental issues (e.g., the rule of law, functioning of democratic state institutions, party system development and functioning). In such cases, sociological or political science narratives push the concept of "civil society" to the background, if they refer to it at all. This happens where the subject of research is a well-consolidated democracy (at least in some of its aspects).

There are times, however, where researchers rediscover the cognitive usefulness of the concept. This takes place, for example, where a stable functioning and reproduction of the democratic order is disrupted and the very existence of democracy is jeopardized. This is also the case when new social movements emerge in the public sphere, contesting a particular part of reality as, for example, ecological movements that have had an impact on political decision makers and made the public opinion sensitive to the consequences of environmental devastation. However, the revival of this concept can be seen most clearly in times of great historical revolutions, leading to a democratic order that emerged after the breakdown of authoritarian systems, as was the case in the 1980s and 1990s.

In different theories of democracy, the analytical usefulness of the concept of civil society varies. Generally, in theories that focus solely on the procedural dimension of democracy, where the problem of civic participation in public life is of lesser importance, the usefulness of this concept is less than in theories focusing on substantial aspects. In the former approach, a civil society is reduced to the so-called third sector, understood as the totality of voluntary nongovernmental associations that articulate various interests in the public sphere. In the latter approach, the usefulness of civil society is greater as the activity of citizens in the public realm determines the vitality of democracy. Examples of such a perspective are the associational democracy model as well as the deliberative democracy model.

Civil Society and Globalization

In the globalized world of today, with the progress of technologies, transport, and, most of all, means of mass communication that cover the entire world, problems that used to be of local impact (humanitarian disasters, epidemics, financial crises) now turn into global issues. These problems are addressed by governments of nation-states, international economic and political organizations (e.g., the United Nations, the World Bank, or the World Trade Organization), as well as private financial or industrial multinational corporations. The global governance of transnational decision-making structures creates the need to ensure civil scrutiny and to build a "global civil society."

However, using the concept of civil society in a global context creates both theoretical and practical difficulties. It is true that some institutions of civil society are becoming globalized. The 1980s were a breakthrough decade in this respect, when some civil initiatives (e.g., human rights and pacifist movements) transcended the borders of the nation-states. They referred to global problems and demanded global solutions as well. Some of those initiatives entered the institutional phase, which resulted in the emergence of nongovernmental civic institutions of global scope (e.g., Greenpeace, Transparency International, Amnesty International). However, the conclusion that we are dealing with the beginnings of a global civil society seems to be premature for a number of reasons.

First, a civil society consists of citizens to whom authorities, legitimated by democratic procedures (e.g., elections), are accountable for their decisions. Global governance cannot be equated with political power defined in this way; because there is no global accountability procedure, this relationship cannot be directly transposed to the global level. Second, there are no rules to legitimize the actions of global civil structures, as a result of which their activity faces claims of usurpation. Third, the emergence of "global civil society" would require a common normative base, on which even contradictory civil initiatives could meet and operate in mutual respect. In a global dimension, there is no such common cultural denominator. Even the concept of basic human rights is not generally recognized and is sometimes interpreted as a product of the Western culture and a tool for its expansion.

The importance of global governance is rising, as is that of global grassroots civil initiatives. However, because there is no "social contract" on the global level, both institutions that execute global governance and global civil organizations suffer from legitimacy deficits. Therefore, their relationship is usually antagonistic. Hence, the concept of civil society may again turn out to be a useful analytical category to describe further developments of globalization.

Edmund Wnuk-Lipinski and
Xymena Bukowska
Institute of Political Studies, University of Warsaw
Warsaw, Poland

See also Accountability; Citizenship; Democracy, Types of; Liberalism; Nongovernmental Organizations (NGOs); Participation; Political Culture; Republicanism; State

Further Readings

Alexander, J. C. (2006). *The civil sphere*. Oxford, UK: Oxford University Press.

Cohen, J. L., & Arato, A. (1992). *Civil society and political theory*. Cambridge: MIT Press.

Colás, A. (2002). *International civil society*. Cambridge, UK: Polity Press.

Edwards, M. (2004). *Civil society*. Cambridge, UK: Polity Press.

Gellner, E. (1994). *Conditions of liberty: Civil society and its rivals*. London: Hamish Hamilton.

Habermas, J. (1989). *The structural transformation of the public sphere*. Cambridge: MIT Press. (Original work published 1962)

Hall, J. A. (Ed.). (1995). *Civil society: Theory, history, comparison*. Cambridge, UK: Cambridge University Press.

Keane, J. (2003). *Global civil society?* Cambridge, UK: Cambridge University Press.

Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America, and post-communist Europe*. Baltimore: Johns Hopkins University Press.

Pérez-Diaz, V. (1993). *The return of civil society: The emergence of democratic Spain*. Cambridge, MA: Harvard University Press.

Civil War

Civil war can be defined as a violent conflict that pits states against one or more organized nonstate actors on their territory. This distinguishes civil wars from interstate conflict (where states fight other states), violent conflicts or riots not involving the state (sometimes labeled intercommunal conflicts), and state repression against individuals who cannot be considered an organized or cohesive group, including genocides or similar violence by nonstate actors, such as terrorism or violent crime. This entry first discusses the various kinds of civil war and reviews the key patterns and trends in such conflicts. It then presents theories on the causes of civil war and the current research agenda.

The above conceptual definition of civil war clearly encompasses many different forms of conflict. Some analysts distinguish between civil wars where insurgents seek territorial secession or autonomy and conflicts where insurgents aim for control of the central government. Conflicts over government control may involve insurgents originating from within the *center* or state apparatus, as in military coups, or challengers may come from the *periphery*, or outside the political establishment. Others separate between *ethnic* civil wars, where the insurgents and individuals in control of the central government have separate ethnic identities, and *revolutionary* conflicts, where insurgents aim for major social transformation. *Colonial* conflicts are sometimes singled out as a set distinct from civil wars on states' core territory. However, even if these suggested categories may be conceptually distinct types and could suggest different causes and dynamics, a given civil war will often combine several elements. For example, insurgencies may be both ethnic and ideologically based, and aims can shift over time from secession for a limited territory to controlling the entire state. Thus, most researchers do not impose such strict a priori distinctions between varieties to analyze these as separate types but instead study civil war as a common class of conflict.

Patterns and Trends in Civil War

Even though the modern state sometimes is defined by its alleged monopoly on violence, armed challenges to state authority are as old as states themselves. Despite numerous historical accounts of civil wars, there are few systematic data sources on civil conflict prior to 1945, since so little is known about the extent of conflicts outside the developed countries before this date. Figure 1 displays the number of ongoing conflicts (top) and new outbreaks (bottom) since 1945, using data from the Uppsala Conflict Data Program at the Department of Peace and Conflict Research, Uppsala University and Centre for the Study of Civil War at the International Peace Research Institute, Oslo (UCDP/PRIO), Armed Conflict Dataset. UCDP/PRIO classifies conflicts claiming more than 25 deaths in a year as interstate conflicts, extrasystemic (or colonial) conflicts, as well as internal conflicts, or as civil wars, including internationalized internal conflicts where other states fight on the side of the government. Figure 1 shows that while there are relatively few interstate wars since 1945, civil wars have been common. Whereas interstate conflicts tend to be short, civil wars often persist for a long time. Research also demonstrates that civil wars are less likely to be settled by formal agreements than interstate wars and much more likely to recur. Many observers saw the initial rise in new outbreaks of civil conflict after the end of the Cold War as evidence that the world would be more turbulent and violent after a period of stable deterrence between the superpowers. However, Figure 1 demonstrates a clear decline in civil war since this peak after the Cold War. The specific causes that may underlie this decline remain disputed, and the number of ongoing civil wars remains high in absolute terms.

Civil wars are generally less severe than interstate wars in terms of the direct battle deaths. However, civil wars have been more frequent and often persistent, and more than 90% of the recorded deaths in battle since the Cold War stem from civil wars. Further, war can have a substantial indirect impact on human welfare beyond the direct loss of life. Studies indicate that countries experiencing civil war see a pronounced fall in their gross domestic product and never recover their earlier growth trajectory. According to Paul Collier et al. (2003), "Civil war is development in reverse." Civil wars also dispel trade and investment and leave large social legacies in unemployed former combatants and displaced individuals. The negative consequences of civil war are not limited only to the countries that experience them; studies find that neighboring countries also suffer a negative economic impact and may be more prone to violence themselves as a consequence of conflict among neighbors.

Causes of Civil War

The literature on the causes of civil war is enormous, and it is impossible to provide an exhaustive review of the many arguments that have been presented here. Contributions have emphasized a variety of social, economic, and political factors. However, theories of civil war, in general, point either to specific motives for why people resort to violence or to specific opportunities that make violence more or less feasible or attractive. This entry examines theories of civil war by the specific clusters of factors emphasized, in an order roughly following the chronology of theory development and the specific events and cases that motivated these arguments.

Economic Grievances, Motives, and Opportunities

Most civil wars take place within relatively poorer societies, and many studies corroborate the link between development and income. Early contributions to the study of violence within societies tended to focus on economic deprivation and grievances as key motives. In particular, Ted Gurr highlighted inequality and how groups may resort to rebellion if they are dissatisfied with their current economic status relative to their aspirations. The literature on nationalist conflicts has emphasized how both relatively poorer and wealthier groups are likely to rebel against the center if they believe that they can do better under independence. Civil wars in Latin American countries were often interpreted within a framework focusing on economic grievances, in the form of either unequal land distribution or high income inequality. However, the empirical evidence linking individual income inequality and conflict is mixed. Older studies focusing on a broad range of political violence often found a positive effect of higher inequality, while newer studies of civil war more specifically tend to find little support. However, one should be cautious in making strong inferences from this, given the poor quality and coverage of cross-national data on individual income inequality.

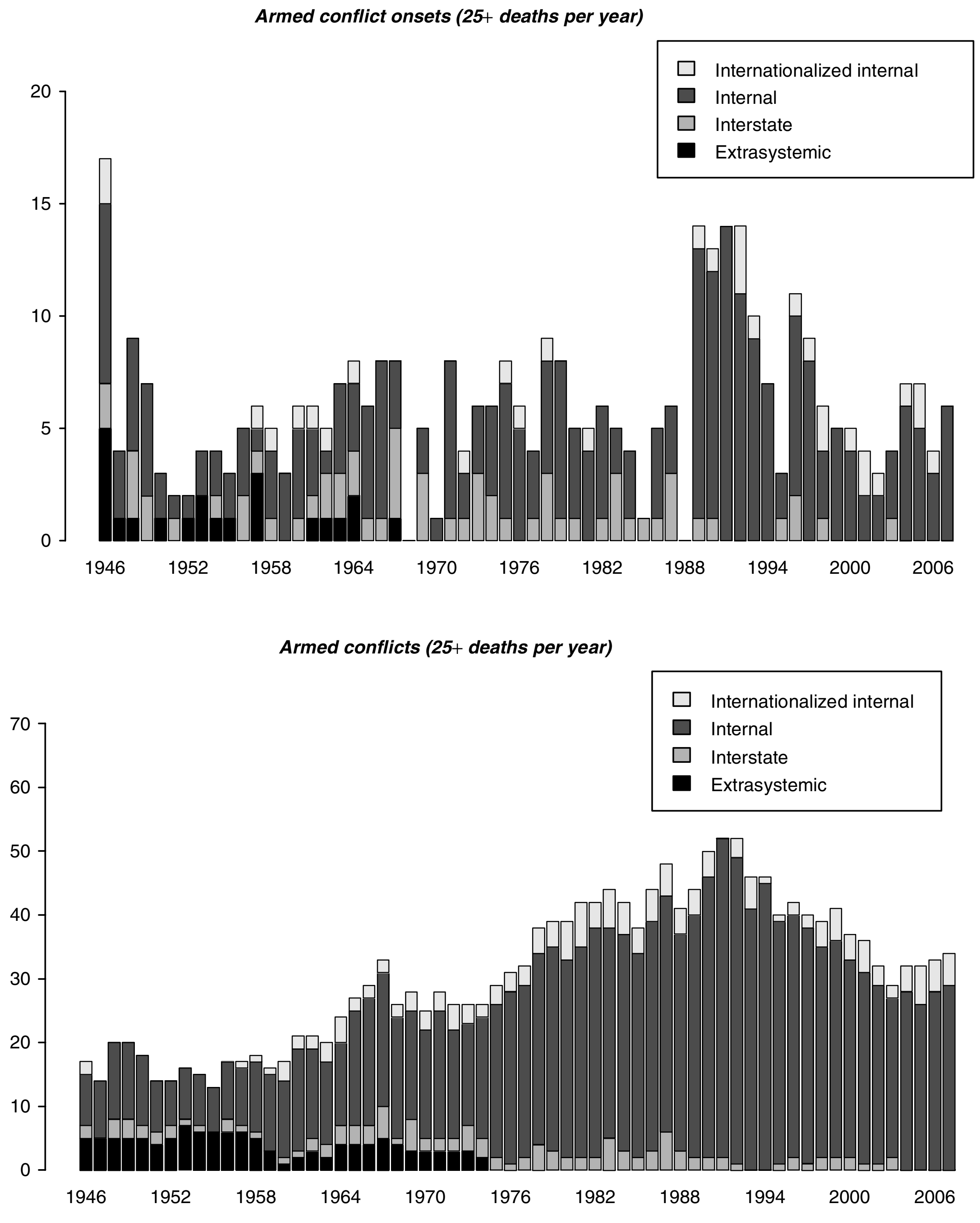


Figure 1 Ongoing Armed Conflicts (*top*) and Onset of New Conflicts (*bottom*)

Source: Author. Data from the UCDP/PRIO Armed Conflict Dataset.

Note: UCDP/PRIO = Uppsala Conflict Data Program at the Department of Peace and Conflict Research, Uppsala University and Centre for the Study of Civil War at the International Peace Research Institute, Oslo.

More recent political economy studies of civil war tend to be very dismissive of the role of grievances. Some researchers have argued that grievances are ubiquitous and that it is more important to focus on variation in the opportunities for violence. Collier and Anke Hoeffler argue that low income makes it easier to mobilize insurgencies, since potential recruits have less to lose in forgone income from normal economic activities and wages for soldering will be much lower than in wealthier societies. James Fearon and David Laitin argue that civil war is primarily a problem of weak states, which in turn is largely determined by economic development. Researchers in this tradition also link mobilization to the role of individual incentives. Opportunities for insurgencies are better when participants can do well out of war, for example, through looting or rents derived from valuable natural resources. Empirical studies lend some support to link between natural resources and a higher risk of civil war. Civil wars in Africa are often taken to support these perspectives.

Political Grievances, Motives, and Opportunities

Political deprivation, such as colonial subordination or lack of political rights, provides another plausible motivation for resort to violence. Many conflicts after 1945 first emerged as groups sought to achieve independence for areas under colonial rule. The Indochina War and the Algerian War of Independence helped mobilize movements in other countries by showing how overwhelmingly more powerful colonial powers could be defeated through sustained violent campaigns. Although overseas colonies eventually received their independence, many ethnically distinct groups within contiguous empires such as Russia or Ethiopia see themselves in similar struggles of national liberation. Researchers have highlighted how violence may arise around such peripheral ethnic minority groups. The violent conflicts in the former Yugoslavia spurred renewed interest in ethnic conflict. Many researchers postulated strong parallels between the security dilemmas in an anarchic international system and relations between ethnic groups, where only full territorial partition could provide stable solutions.

Much of the subsequent empirical literature has emphasized that there is little evidence that ethnically more diverse countries are generally more prone to conflict. However, looking only at the number or relative size of ethnic groups disregards their political status and the extent to which ethnic groups are systematically excluded from political power or discriminated against by the state. Studies looking at ethnic exclusion find stronger evidence for a relationship to violence. Ethnically diverse countries are not necessarily more prone to conflict if they have inclusive institutions or grant autonomy rights, and control of the state or access to power does not always follow directly from the relative size of ethnic groups. Many ethnic civil wars see minority groups controlling political institutions and excluding larger groups from power, such as the Amhara in Ethiopia until their defeat by a coalition of other groups in 1991. Political, economic, and ethnic grievances are often linked. Studies that focus on inequalities that follow group lines rather than individual income inequalities find a stronger relationship to conflict.

Struggles for broader political rights in autocratic systems provide another context where violence may occur. Political democracy provides many avenues for actors to express dissent through nonviolent political means. Autocratic regimes typically deny citizens room for political activities and often resort to severe repression of protest, which in turn may motivate resort to arms. Protests against autocratic or exclusionary regimes have often turned violent, sometimes leading to sustained campaigns, as, for example, in South Africa. Claims for greater political rights and freedom are clearly important elements of the rhetoric and call for mobilization of many insurgent movements, even if these do not necessarily move to implement democratic institutions if successful in achieving power.

Much of civil war research has focused on the role of political institutions in structuring the opportunities for violence rather than the potential accommodative effect of greater democracy. Many have argued that although autocratic institutions provide fewer avenues for nonviolent political activities and protest, autocratic regimes are often sufficiently repressive to successfully deter dissent. By contrast, anocracies, or regimes combining autocratic and democratic features, are seen as the most prone to see violent conflict by combining the lack of political freedom motivating violence with the sufficient opportunities afforded by a less repressive regime. Many studies have found evidence for such an inverted U–shaped relationship between democracy and civil war. However, there is no consensus on the underlying mechanisms (e.g., reforms in autocracies may be a response to conflict potential rather than a prior cause). Some hold these relationships to be in part artifacts resulting from the definition of empirical democracy measures.

Opportunity Structures and Dynamic Events

Most of the theories discussed above emphasize structural factors that rarely change or change only slowly over time. Such persistent structural features do not provide clear explanations for why civil wars break out at specific times and not others. Research on social movements suggests that certain events can create "political opportunity structures" that afford groups better prospects for extracting concessions from the state or center. This may include demonstrations of state weakness, conflict between elites, or events that make it easier for groups to mobilize, for example, by bringing groups together or indicating focal points for organizing protests. The concept of opportunity structures has so far not had much direct impact on studies on civil conflict, but there are many existing arguments and findings in civil war research that can be interpreted within this framework. Regime change and other signals of weakened state authority can increase the perceived chances of success or extracting concessions from a government. Studies have shown that economic crises and natural disasters can increase the risk of conflict. This is consistent with the idea that crises and emergencies can help provide a setting for rallying protest against the government. For example, the 1973 earthquake in Nicaragua, and the massive corruption and lack of subsequent reconstruction, generated widespread disillusionment and helped a long-standing Marxist insurgency dramatically increase recruitment.

International Dimensions of Civil War

Most research on civil war has assumed that since civil wars are "internal" conflicts rather than conflicts between states, their main causes must also be domestic or located within state boundaries. However, factors outside individual countries can play an important role in the outbreak of conflicts, as well as how they evolve. Many actors in civil wars are not necessarily confined within individual countries. Ethnic groups often span international boundaries, and transnational kin frequently participate in or provide support for insurgencies in other states. The status of international borders generates different constraints and opportunities for governments and rebels. Borders are, in a technical sense, just lines in the sand and are often not difficult to cross from a purely military perspective. However, the fact that borders formally delineate state sovereignty makes it more difficult for governments to violate the sovereignty of other countries, while such constraints are less relevant for rebels. Governments risk retaliation from neighboring countries from territorial incursions and face difficulties in targeting transnational support. This in turn means that rebels can have a logistic advantage in operating out of extraterritorial bases, and transnational rebel movements can be more difficult for governments to deter or defeat. The presence of conflict in another state can help facilitate violent mobilization, either through emulation of successful rebellions or through the direct imports of arms and combatants. Finally, civil wars are often closely linked to interstate war. Poor relations between states may motivate governments to support insurgencies in rival countries, and civil wars may in turn promote military conflict between states, for example, as a result of conflict over border violations, alleged support for insurgents, or conflicts over the externalities generated by conflicts. Western Africa in the 1990s, for example, provides many examples of governments supporting insurgencies in neighboring countries and retaliating against alleged support and border violations. The fact that conflicts are not necessarily limited to single countries and may involve participation by other states in various forms demonstrates how a strict dichotomy between civil and interstate wars often may be untenable.

Civil War Termination and Duration

Although most research focuses on accounting for the original outbreak of civil war, there has also been growing interests in understanding the prospects for conflict termination and why some conflicts are so persistent. Many researchers argued that there must be some symmetry between the original causes of civil war and the factors that lead to their eventual termination. From this perspective, efforts to foster conflict settlement should focus on addressing the issues that gave rise to the conflict. For example, civil wars would be likely to end earlier in more developed countries, where the opportunity costs are higher and states are stronger. Events such as political reform and increasing accommodation could help promote the end of conflicts. However, other researchers argue that the factors that make conflicts endure can be quite different from those influencing the initial outbreak of conflict. While interstate wars tend to be relatively short and typically end quickly in some kind of formal settlement once the relative strength of the actors becomes clear, civil wars often persist for a long time, as in the case of Southern Sudan or Angola. A number of explanations have been proposed to account for this paradox. Some emphasize how civil wars tend to be low-intensity conflicts fought in the periphery, where the features that facilitate insurgency make it difficult for governments to conclusively defeat rebels. Others point to how insurgents that can do well during war through looting and control over valuable resources may have little interest in seeking an end to conflict. Other research emphasizes how conflict termination involves severe problems of credible commitment. Even if it may be relatively easy to reach consensus on the terms of an agreement in principle, carrying through with an agreement may be difficult since parties have incentives to make promises that they later renege on. For rebels, laying down arms is a risky choice, since they will be left less able to defend themselves. Moreover, it can be difficult for governments to ascertain whether rebels will uphold their end of agreements, for example, through decommissioning or controlling violent breakaway factions in the aftermath of agreements. Research on credible commitment problems after civil war argues that the success of agreements often hinges on whether third parties can serve as external enforcers. New research on peacekeeping and conflict management suggests that involvement by international organizations such as the United Nations (UN) can increase the prospects for settling conflict and preventing wars. However, many interventions in civil conflicts have clearly been less than fully successful. There is little consensus on how best the UN may contribute or to what extent the characteristics of conflicts or initial prospects for settlements determine whether external actors are willing to be involved in the first place.

Disaggregation in Civil War Research

Civil war has gone from being a peripheral area to a major focus in the study of conflict. Moreover, the field shows clear positive synergies from the interaction between theory development and research design. Older studies of civil war were often quite descriptive, focused on single cases, and frequently reluctant to consider general theories of civil war. The renewed interest in civil wars in the 1990s recognized the need to take into account nonconflict cases and turned to cross-national studies, comparing the various characteristics believed to make countries more prone to conflict across countries and its relationship to civil war. However, recent research has highlighted how the proxy measures used in cross-country comparative studies are often quite far removed from the key theoretical concepts. Moreover, looking at states and countries at large often downplays the nonstate actors in civil wars and ignores the often considerable variation within countries. Conflicts tend to be localized and often involve actors and regions that are markedly different from national averages, as, for example, the Chechen conflict in Russia. Recent contributions, accompanied by many new data developments, often turn to greater disaggregation within countries and a more specific focus on actors to more carefully evaluate theories of civil war. A number of studies have looked at the specific conflictual events within individual countries over time to better understand the microlevel interactions between actors in conflict. Some studies have looked at dyads of peripheral groups and the center to take into account how the attributes of nonstate actors may influence the risk of conflict. Other studies have looked at the characteristics of specific rebel groups to better understand capacity or opportunities for conflict or used individual-level data on participants in insurgencies to examine motivations. Some researchers consider geographically disaggregated data, either on the characteristics of the places where conflict occurs or using smaller units within countries, for a more detailed resolution of how social, economic, and political factors may be related to conflict. These studies indicate that the characteristics of conflict zones tend to be quite different from national aggregates and averages. Moreover, the conclusions from country-level studies on who participates in insurgencies and their motivations are often not supported by more direct evidence. Such innovations in the study of civil conflict attest to the vibrancy of this research area. Moreover, they underscore the important relationship research design and the conclusions that we reach, and how theory must inform the former. Although existing research has clearly generated many useful insights about civil war, this review illustrates the considerable changes in the conventional wisdom on civil war over time. Advances in theory development and research are likely to continue to spur further changes.

Kristian Skrede Gleditsch
University of Essex
Colchester, United Kingdom

See also Authoritarian Regimes; Conflicts; Democratization; Dissatisfaction, Political; Ethnicity; Inequality, Economic; Mobilization, Political; Nationalism; Nonstate Actors; Peace; Revolution; Secession; State Failure; Violence; War and Peace; Warlords

Further Readings

Blattman, C., & Miguel, E. (2010). Civil war. *Journal of Economic Literature, 48*(1), 3–57.

Buhaug, H., Cederman, L.-E., & Rød, J. K. (2008). Disaggregating ethno-nationalist civil wars: A dyadic test of exclusion theory. *International Organization, 62*(3), 531–551.

Cederman, L. E., & Gleditsch, K. S. (Guest Eds.). (2009). Introduction to special issue on "Disaggregating Civil War." *Journal of Conflict Resolution, 53*(4), 487–495.

Collier, P., Elliott, L., Hegre, H., Hoeffler, A., Reynal-Querol, M., & Sambanis, N. (2003). *Breaking the conflict trap: Civil war and development policy.* Oxford, UK: Oxford University Press.

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review, 97*(1), 75–90.

Gleditsch, K. S. (2007). Transnational dimensions of civil war. *Journal of Peace Research, 44*(3), 293–309.

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed conflict 1946–2001: A new dataset. *Journal of Peace Research, 39*(5), 615–637.

Gurr, T. R. (1970). *Why men rebel.* Princeton, NJ: Princeton University Press.

Kalyvas, S. N. (2006). *The logic of violence in civil war.* Cambridge, UK: Cambridge University Press.

Salehyan, I. (2009). *Rebels without borders: State boundaries, transnational opposition, and civil conflict.* Ithaca, NY: Cornell University Press.

Class, Social

The concept of class is one of the most important and most controversial in social sciences. It is important because classes have been traditionally considered a key to the study of social inequalities and social and political behavior. It is highly controversial because different definitions and contrasting perspectives on the role and impact of class have been proposed. In everyday language, class refers to social groupings based on similar occupational conditions that have the same access to economic, social, and political resources. It is also assumed that experiencing analogous social inequalities favors the formation of a specific cultural identity and influences social and political behavior. However, in scientific language, classes are to be considered a "modern" phenomenon. This means that whereas social inequalities have always been present in human history, the specific inequalities based on class require two conditions: the equality of citizens before the law and the development of a market economy. These conditions occur after the end of the 18th century with the advent of modern society, shaped by democratic revolution and growth of the market economy and industrialization.

Therefore, classes can be defined as social groupings that are based on similar occupational conditions in societies where a market economy prevails and the formal equality of citizens before the law is recognized. People belonging to the same class de facto experience an unequal access to economic, social, and political resources depending on the jobs they occupy, despite the formal equality of rights. However, they can change their class through social mobility. This possibility clearly distinguishes classes from other social groupings, such as religious castes or medieval estates, for two main reasons. First, in the case of castes or estates, inequalities are established by traditional laws. Second, the class position of people cannot be changed, because social mobility is not allowed (as for castes) or is subject to special conditions defined by law (as in the case of medieval estates). In modern societies, mobility is a normal condition, although classes presuppose a certain degree of enclosure.

While this definition of class is widely accepted in the literature, sharp differences have emerged along two dimensions. First, opinions diverge on the economic base to be used to analyze the class structure. Second, strong disagreement persists on the cultural, social, and political consequences of classes. On the one hand, this is related to the specific mechanisms that trigger the transformation of classes from mere aggregates of people into social groups: collective actors that recognize common values and interests and organize themselves to pursue common goals. On the other hand, scholars differ in the evaluation of the impact, over time, of classes on social behavior (lifestyle, consumption) and political behavior (participation, voting).

In the next section, the main approaches to the origins and consequences of social classes are discussed. First, the perspectives of classical authors such as Karl Marx, Max Weber, and Émile Durkheim are contrasted. The influence of these authors on the ensuing research is extensive. Contemporary studies are still affected by their different views. After discussing the most important proposals for reframing class structure in the light of economic and political changes, attention is devoted to debate over the past decade on the decline of social classes: To what extent do social classes remain important as an instrument for understanding contemporary societies?

Contrasting Views in the Classics

Among the classicists, Marx is the author who gave greatest emphasis to social classes. The struggle between classes plays a fundamental role in the historical development of human societies. Classes are social groupings that share a specific position in the control of the means of production. Each historical type of society is characterized by a dominant mode of production and specific classes. The capitalist mode of production is based on the private property of the means of production and on the role of the market in regulating productive activities and the distribution of incomes. Capitalist societies were therefore increasingly shaped by the capitalist mode of production with its two main classes: (1) the owners of the means of production (capitalist bourgeoisie) and (2) the workers, who are employed by the former and receive a wage. According to Marx, there was an objective conflict of interest between these two classes, because the industrial bourgeoisie could maintain, and strengthen, its position of economic, social, and political privilege as far as it was able to exploit the workers and extract from their activities a surplus value as determinant of profit. Over time, however, the growing crises of the capitalist economy would worsen the conditions of the working class and would encourage its social and political organization, which, in turn, led to a revolutionary change in the mode of production and the advent of a socialist society.

Marx was aware that the class structure of capitalist societies was more complex. At the same time, he recognized that the making of the working class as a historical actor did not depend on economic factors alone. He frequently discussed the sociocultural and political factors that could affect the formation of class consciousness (as he puts it, the passage from a "class in itself" to a "class for itself"). However, over time the extension of the capitalist mode of production would lead to the increasing polarization of bourgeoisie and workers and to an overwhelming trend toward conflict. The history of the past 2 centuries has shown the importance of classes and class struggle, but it has also demonstrated that Marx overestimated the role of class conflict and overlooked the autonomous role of social and political processes and institutions.

Weber was convinced of the important role played by class and class conflict. However, he saw class structure in a more flexible way, which allowed him to account for the increasing differentiation introduced by the development of a modern capitalist economy. This was done by defining classes as social groupings based on a common position in the market rather than on the control of means of production, as Marx asserted. Therefore, a class is composed of people who have similar life chances, affected by the income they can obtain in the market. This common power in the market ("class situation") is determined not only by the ownership of the means of production but also by wealth, education, and professional skills.

Another difference concerns the passage from a common class situation to the formation of self-aware social groups that organize themselves to pursue common interests. While Marx was convinced that the economic dynamic of capitalism would lead, in the end, to class consciousness and collective action, Weber was much more skeptical about this possibility. He thought that people within the same class situation would share life chances, social inequalities, styles of consumption, social and even political behavior, but they would usually remain aggregates of individuals rather than becoming self-aware social groups based on social interaction. In this respect, the German sociologist distinguished between mass action and community action. A common class situation usually brings about mass action. But the passage to community action requires certain noneconomic conditions: Status groups (religious, ethnic groups) must be absent or marginal; a large number of people must be in the same class situation; because they are concentrated in large firms and urban neighborhoods, it is relatively easy to organize them; and intellectuals will play an important role in the diffusion of a class ideology and in the organization of the people involved. Despite these clearcut differences in their view of classes and class conflict, both Marx and Weber recognized the importance of social inequalities as a determinant of social behavior and source of conflict in modern societies.

A different perspective was adopted by another founding father of sociology, Durkheim. Writing at the end of the 19th century, he emphasized the increasing division of labor as the main feature of modern society. A fundamental requisite of this society, based on the high differentiation and integration of specialized roles, concerns the recruitment and reward of individuals. It is necessary that each individual be called to fulfill the function he or she will perform best and receive an adequate reward for his or her efforts. In other words, people must be motivated to perform different roles thanks to structured social inequalities. This is the main function of stratification in the social system according to Talcott Parsons, the most influential figure in the sociology of the 1950s, who developed the functionalist premises of Durkheim's work. Parsons and other scholars of the functionalist approach, such as Kingsley Davis and Wilbert Moore (1945), gave a more optimistic account of the role of social stratification in the stability of society. This trend—which was also influenced by the economic and social climate of the postwar period—fostered the growth of descriptive research on social stratification, especially in the United States. Social strata were constructed by reputational analysis (asking people to rank others) or by a mix of reputational evaluations and objective indicators, such as wealth, income, and education. Usually, as in the well-known work by Lloyd Warner (1963), these studies presented a classification based on various status groups (upper classes, middle classes, and lower classes).

Reframing Class Analysis

In the 1950s and 1960s, the stratification approach largely prevailed in the study of social classes. This situation changed in the 1970s with the development of new empirical research based on occupational classes. In this strand of literature, there is a shared critique of the descriptive features of stratification studies and an emphasis on the "relational" attributes of classes. This means classes are to be defined on the basis of their position in the market and in the productive process, as in Weber and Marx, and therefore in relation to the other classes with which they compete for greater resources. However, the origins and aims of the new studies are clearly different. The component most influenced by Marx and the Marxist tradition is more interested in the analysis of the middle classes and in its consequences for the proletarianization of work, originally foreseen by Marx. The other component is closer to the Weberian perspective and looks at the empirical study of classes as the necessary basis for a more adequate analysis of social mobility.

The followers of the Marxist perspective tried to account for both the persistence of the petty bourgeoisie and the unexpected (in Marxist terms) growth of the new, dependent, middle classes (managers, supervisors, technicians). In designing an empirical map of social classes, Erik Wright tried to solve this problem through the concept of "contradictory class locations." In his studies, he distinguished between the ownership and control of the means of production. This allows him to explain the new middle-class positions as based on various degrees of control over the concrete use of the productive means without formal ownership. As a consequence, top and middle-level managers, technicians, and supervisors may also experience variable degrees of autonomy in performing their role.

In the 1980s and 1990s, another influential class map was provided by John Goldthorpe. He distinguished between a labor contract and a service contract. Employment regulated by a labor contract entails a short-term and specific exchange of money for work, while a service contract involves a delegation of authority, specialized knowledge, and a certain degree of autonomy on the part of employees. On this basis, a class schema for empirical analysis and comparisons was worked out. It included seven main classes: (1) service class (managers, professionals, higher-grade technicians), (2) less skilled white-collar workers, (3) petty bourgeoisie (artisans, small proprietors), (4) skilled manual workers, (5) unskilled manual workers, (6) small farmers, and (7) farm laborers. In practice, this schema is similar to the map designed by Wright. Both account for the growth of the new middle classes and an increasing differentiation among manual workers. Both of them can be used in diachronic and cross-country comparisons. For example, over time the class structure of advanced countries shows a decline of the old middle classes (the petty bourgeoisie); an increase in professionals, technicians, and white-collar workers (the new, dependent middle classes); and a growing differentiation among manual workers. Cross-country comparisons may help analyze the differences among advanced countries, which are related not only to the productive model and organization but also to the role of the state and the extension of nonmarket forms of coordination ("varieties of capitalism"). Interesting insights may also come from comparisons between advanced and developing countries. However, the main application of the Goldthorpe schema has been in the comparative study of social mobility. In this respect, robust research findings have been produced on the strong influence of classes on social mobility. These findings do not confirm the expectations of the liberal theory that industrial societies tend to become more mobile and more open. Rather, mobility rates (in particular relative rates) show a high degree of temporal stability. The chances for the children of less privileged classes of moving to the upper positions of the class structure remain low, without revealing significant differences among countries.

In spite of their different origins and goals, both the Marxist and Weberian maps achieved similar results and shed light on the class-based social inequalities typical of advanced capitalist societies until the 1970s. However, in the following period, there was a double change. On the one hand, the crisis of the "Fordist" model of production brought about a decline of large firms and their organization of labor. New forms of productive organization developed, based on flexible specialization and networks linking firms. On the other hand, welfare systems had to be restructured to deal with increasing costs. As a consequence, important changes affected the whole class structure. New maps had to be designed. "Postindustrial" societies are characterized by the growing role of the service sector. The traditional axis of stratification based on industrialism has shrunk, both because of the new technologies that reduce the use of labor and because of increasing trends to outsource manufacturing activities to developing countries. In advanced capitalist societies, new activities have developed in business services (finance, legal and accounting services, software programming), in social services (health, education, and welfare services), and in consumer service (restaurants, laundries, services related to leisure). Therefore, as Gøsta Esping-Andersen has pointed out, a new axis of stratification has emerged. This requires that new groupings be more clearly distinguished within the middle classes and the workers. The roles of professionals and scientists, together with that of technicians and semiprofessionals (schoolteachers, social workers, technical designers), become more important in the middle classes and have to be differentiated from those of managers and supervisors. Among the workers, it is necessary to distinguish new figures such as skilled service workers (police officers, hairdressers, etc.) and unskilled service workers or service proletariat (cleaners, wait staff, shop assistants, etc.). The new groupings tend to experience specific forms of career regimes and life chances and therefore are involved in processes of class closure that differentiate them from the traditional classes based on the Fordist industrial model.

Persistence or Decline?

Two macrochanges have influenced the debate on social classes in the past decades: (1) postindustrialism and (2) globalization. According to many scholars, the crisis of Fordism, the emergence of a service-based economy in advanced capitalist countries, and the increasing delocalization of manufacturing activities toward the developing countries have brought about a decline in social classes. The basic arguments in support of this thesis refer to two main trends: On the one hand, the increasing fragmentation of class structure and the growing individualization of life chances, and on the other hand, the weakening of class-based attitudes and behavior. For example, according to Ulrich Beck (1992), social inequalities tend to become "classless," whereas Anthony Giddens pointed out that—as a consequence of fragmentation and globalization—classes are no longer experienced as a significant source of collective identity.

Those who share the idea of a class decline usually refer to both dimensions: the loss of influence over social inequalities and social behavior. In contrast to this thesis, other scholars strongly criticize this view of a decline. However, they mainly focus on the relationship between classes and social inequalities. Goldthorpe provided a clear synthesis of the latter position, which makes wide reference to empirical findings. He pointed to the more detailed arguments advanced in support of the decline: insecurity and mobility. The theorists of decline assume that within the context of globalization, the threat of unemployment is not confined mainly to the members of the less advantaged classes but tends to become pervasive. As for mobility, globalization and postindustrial forms of economic organization are supposed to break up the continuity of working careers and of membership in the traditional occupational classes. The growth of flexible and nonstandard work and the emergence of new, more fragmented work trajectories reduce the significance of class as a lifetime experience. However, in contrast to these expectations, empirical findings show the persistence of class as the main determinant of inequalities. Thus, it is true that economic changes have brought about increasing job insecurity and a growth of nonstandard employment, but the chances of losing their jobs remain much higher for skilled and especially nonskilled workers than for professional, administrative, and managerial posts. In addition, empirical research demonstrates that inequalities in important spheres such as income, health, and access to education remain strongly associated with the class of people, defined according to the job they hold. As for social mobility, there is clear evidence that the overall association between class origin and destination is still characterized by high temporal stability and, therefore, despite changes in the organization of work, classes continue to provide lifetime experiences.

In the debate on class decline, particular attention has been devoted to the influence of class on voting and political participation. According to the theories that point to the weakening of class experience, class politics gives way to new forms of political behavior influenced more by cultural factors, lifestyles, and elective choices. One of the most powerful factors that feed this trend is the dramatic reduction in size of the traditional working class and the growth of service workers and the new middle classes. As a consequence, class cleavages become a less important basis for social identity. Their place is increasingly taken by other divisions, such as race and ethnicity, gender and sexuality. Many scholars also share Ronald Inglehart's (1990) thesis of the new "postmaterialism" in the politics of the advanced societies. This view sees politics no longer as influenced by the search for economic well-being and the defense of material interests but as increasingly affected by a personal search for autonomy and self-fulfillment and by a greater cultural interest in the quality of life.

As in the case of social inequality, scholars mainly involved in empirical research have rejected the thesis of a declining class influence on politics. They have pointed to findings that clearly question the hypothesis of a class dealignment—a reduction in the level of class voting—controlling for changes in the relative size of classes and the overall popularity of parties: Old relations between classes and parties—such as those of workers and left-wing parties—still hold. It has also been noted that a possible realignment will usually reveal new links between class and voting patterns rather than their disappearance. Thus, in concluding his accurate comparative study, Geoffrey Evans pointed out that the theory of a generalized decline in the class basis of voting is wrong.

However, in assessing the influence of class, a more clear-cut distinction has to be made between the impact on voting patterns and party affiliation and the strength of social classes as political actors, that is, as self-aware collective subjects, capable of pursuing their own interests in the economic field as well as in the political arena. Empirical findings show the persistence of this impact on voting and party affiliations, but the trend does not contradict the weakening of class on another dimension: the formation of self-aware social groups. There are clear signs of a decline in this respect, especially with reference to the role of the working class in advanced countries. Its decline in size, the fragmentation of working conditions, and the growth of the service economy have substantially weakened the working class as a social and political actor. This trend is confirmed—with variations from one country to the next—by a decline in trade union membership and union strength and by the changes that have affected left-wing political parties in their attempt to adjust to the transformations in class structure. The opening up of income differentials, at the expense of the less privileged classes, the restructuring of welfare systems, and the overall weakening of social protection are clear indicators of this decline.

Therefore, it may be said that class still clearly influences social inequalities and life chances (income, health, education, social mobility), but when it comes to action its impact is weaker. To return to Weber's distinction between mass action and community action, one can maintain that the influence of class on social and political behavior (i.e., on some aspects of consumption and on voting) usually takes the form of mass action involving aggregates of individuals. Over time, however, the impact of class on the formation of self-aware social and political groups through community action has declined.

Carlo Trigilia
University of Florence
Florence, Italy

See also Cleavages, Social and Political; Conflicts; Electoral Behavior

Further Readings

Beck, U. (1992). *Risk society: Towards a new modernity*. London: Sage.

Dahrendorf, R. (1959). *Class and class conflict in industrial society*. Stanford, CA: Stanford University Press.

Davis, K., & Moore, W. E. (1945). Some principles of stratification. *American Sociological Review, 10,* 242–249.

Esping-Andersen, G. (Ed.). (1993). *Changing classes: Stratification and mobility in post-industrial societies.* London: Sage.

Evans, G. (Ed.). (1999). *The end of class politics? Class voting in comparative context*. Oxford, UK: Oxford University Press.

Giddens, A. (1973). *The class structure of the advanced societies*. London: Hutchinson.

Giddens, A. (1994). *Beyond left and right: The future of radical politics*. Stanford, CA: Stanford University Press.

Goldthorpe, J. H. (2002). Globalisation and social class. *West European Politics, 25,* 1–28.

Hall, P., & Soskice, D. (Eds.). (2001). *Varieties of capitalism: The institutional foundations of comparative advantage.* Oxford, UK: Oxford University Press.

Inglehart, R. (1990). *Culture shift in advanced industrial society*. Princeton, NJ: Princeton University Press.

Warner, L. W. (1963). *Yankee city*. New Haven, CT: Yale University Press.

Wright, E. O. (1985). *Classes*. London: Verso.

Cleavages, Social and Political

The term *cleavage* identifies social and political divisions characterized by a close connection between individuals' positioning in the social stratification system, their beliefs and normative orientations, and their behavioral patterns. This close connection contributes to the resilience and stability of cleavages over time. In this entry, the constitutive elements of this concept are discussed and brought into a coherent overall framework that helps illuminate the political relevance of such cleavages.

The Concept

Empirical evidence about individual actors in politics concerns these three distinct dimensions: social structural, attitudinal, and behavioral. Individuals can be regrouped in social aggregates on the basis of gender, residence, education, property, ethnic group, religion, competence in political matters, available social relationships, and so on. Scholars regroup individuals in this way because some theory or previous research has led them to believe that these social aggregates have an impact on the other two kinds of information we can collect about the same individuals: (1) their subjective perceptions and judgments about the world of politics and (2) their actual political behavior.

Individuals in fact have opinions, preferences, and value judgments about a large variety of political issues, institutions, and actors, as well as normative views on good government, a just society, welfare and redistribution, race and ethnicity, religion, and so on. Some of these beliefs and values are superficial and unstable and change rapidly and frequently, while others are profound and resilient and more impermeable to environmental changes. In some cases, these various beliefs are unrelated to each other, while in other cases, they are strongly correlated and constitute clusters or systems of beliefs.

We also can collect information about what individuals do—their behavior. People establish wide or narrow social relationships (they marry, do business, establish friendships) in their environment. Some individuals invest heavily in culture and education, while others have different consumption patterns. Some cooperate easily with others while others withdraw. (See Table 1.) People vote or abstain, join parties or interest groups, rebel or acquiesce, and, more generally, engage or do not engage in a vast repertory of political acts.

Most sociopolitical research is based on the attempt to investigate relationships between social positions and aggregates, beliefs systems and value orientations, and behavioral patterns. While it is clear that objective social positions are derived from theories of social stratification criteria that are independent of the normative orientations and the behavioral patterns of the individuals, the relation between beliefs and behaviors is more intimate and the direction of influence more difficult to ascertain. These three sources of information should, therefore, remain conceptually distinct as much as possible.

The literature frequently associates these three sources of information about individuals with different kinds of "cleavages."' Objective social positions are often thought to be the basis for "social" cleavages (sometimes called "economic" cleavages);

Table 1 Types of Information About Individuals

"Objective" social positions	Attribution of individuals to "relevant" social categories as defined by the observer
Beliefs and value orientations	Subjective orientations toward political and social objects, institutions, and actors
Behaviors	Repertory of socially and politically relevant behaviors

beliefs and values are considered to be the source of "cultural" cleavages and also of "political cleavages"; and political behavior is often indicated as an expression of political cleavages. The term *cleavage* is therefore frequently accompanied by qualifying adjectives such as social, political, economic, attitudinal, behavioral, cultural, ideological, structural, segmental, institutionalized, and so on. This generates considerable theoretical confusion and has led to a soft use of the term *cleavage* as a general concept that is largely synonymous with the broad notion of divisions, oppositions, or conflicts.

In fact, the historical divisions to which the term *cleavage* originally was applied by Seymour Martin Lipset and Stein Rokkan in reference to party system formation (center–periphery, urban–rural, class, religion, etc.) included socioeconomic bases; attitudinal, ideological, or, more generally, normative bases; and behavioral or organizational bases. None of the historical cleavages was ever purely "economic" or "demographic," merely attitudinal or ideological, or exclusively behavioral. On the contrary, they represented a peculiar combination of these divisions, with a tendency to overlap and to reinforce each other. Indeed, the peculiarity of the concept of "cleavage" is its capacity to bridge social with ideological or normative and with behavioral/organizational divisions. In short, the theoretical connotation of the concept of cleavage refers to the combination of *interest orientations* rooted in social structure, *cultural/ideological orientations* rooted in the normative system, and *behavioral patterns* expressed in organizational membership and action.

To properly define and identify cleavages, therefore, it is necessary to locate them in relation to the various kinds of "divides" that exist and are continuously generated and redefined in our societies. Cleavages then emerge as a peculiar combination of different types of divisions. This is achieved by first discussing separately the sources of divisions rooted in the (a) social stratification system, (b) normative system, and (c) behavioral system.

The Social Stratification System

Pure divisions of interest rooted in social stratification, based on property, market positions, education, religious affiliation, and so on, are unlikely to constitute a sufficient base for the structuring of cleavages unless they tend to overlap and reinforce each other. The higher the frequency and the degree of overlapping membership in different social contexts, the more unlikely it is that any such social context will be dominant.

Individuals play a range of different social roles, some of which imply competing and contradictory tasks, activities, expectations, and values. Individuals' group memberships (in the family, employment, friendship, communication, and leisure activities) and relative social positions (with respect to education, property, income, class, religion, ethnic group, etc.) in modern societies are usually highly differentiated, and the chances are low that many individuals will share the same combination of stratification positions and group membership. At the extreme, however, there may be only a few socially relevant categories, which are characterized by mutually exclusive membership, as in a strongly stratified caste society or a society with strong ethnic, religious, or similar divisions. Then, a high proportion of agents act exclusively in one category or social group. This second situation generates and exacerbates potential divisions among self-containing groups whose nonoverlapping but reinforcing membership facilitates common values and perceptions but makes an exchange of information, experiences, and so on across groups much more difficult. This accentuates group conflicts, particularly because closed groups will tend to claim the complete and total involvement of their members toward some collective aspiration or design.

In low-mobility and high-status contexts, the social boundaries of social groups may be so strong and so institutionally reinforced that exit options at the individual level are basically unthinkable. This reinforces cultural solidarity in the group and reduces individualistic options in favor of collective choices. The more the empirical evidence approaches this situation, the greater the likelihood that the similarity of social positions and interests will give birth to strong and cohesive normative visions of the group's role in society—that is, to a code of collective identity.

The Normative System

At the core of all codes of collective identity is a distinction between "we" and "the others." A typology of the symbolic codes of collective identity distinguishes between codes of primordiality, civility, and culture.

Primordial elements of collective identity focus on gender, generation, kinship, ethnicity, and race as bases for constructing and reinforcing the boundaries between inside and outside. This code of closure makes reference to original and unchangeable distinctions that are by definition exempted from communication, reflexivity, and exchange among the members of the community. Such distinctions cannot be changed by voluntary action; they relate to "nature" and as such provide a firm and stable basis beyond the realm of voluntary actions and thus do not allow for shifting involvements. In other words, primordial types of codes basically "naturalize" the constitutive boundary between insiders and outsiders, and they tend to be exempted from social definition and alteration. Therefore, primordial boundaries are very difficult to change, and complex rituals to cross these boundaries actually reinforce and reaffirm them.

Identity boundaries built on the *civility code* of closure and exclusion are constructed on the basis of familiarity with implicit rules of conduct, traditions, and social routines that define and demarcate the boundary of the community. Tradition and collective civic rituals are important (the myths of the founder; particular persons, places, and historical events such as revolutions; migrations, etc.). Very often these are embedded in the constitutional practices of the public arena. There are some civic rituals but no peculiar rituals of initiation, commitment, or confession to demarcate the boundaries. The stranger is primarily "unfamiliar." Civility codes allow a stronger differentiation between the "public" and "private" spheres, between polity and the market, and between the entitlements to public goods, which are restricted to members of the community, and access to commodities exchanged in the market, which are open to everyone on the basis of material rewards and instrumental calculation and which, therefore, are identity free.

The third code of collective identity—*the cultural code*—is based on the boundary between "us" and the "other" on a particular relation of the collective subject to some fundamental values. The collectivity so defined on the basis of values related to God, reason, progress, rationality, or to some sort of definition of the ideal mundane or heavenly world. Boundaries between members and nonmembers can be crossed by communication, education, and conversion, and in principle everyone is invited to do so. The cultural code usually represents a rupture between past and present, in strong opposition to primordiality codes that always overemphasize the past as the source of legitimacy. The openness of the boundary is compensated by stratified access to the center by rituals of initiation.

Primordial, civic, and cultural codes of identity can be ranged in terms of difficulty in crossing the boundary of the identified community; that is, to exit and to enter the social relationship. We go from the exclusive codes, based on primordial traits, to the inclusive and cultural codes, while civility codes are located in an intermediate position between these two extremes.

The level of identity, solidarity, and loyalty generated by different codes entails consequences for the allocation of resources and for the definition of the entitlements of the members of the collectivity as distinct from those of outsiders. Groups' identities differ with respect to what entitlements are distributed in the name of the collectivity to its members and the mode of distributing those entitlements (hierarchical vs. egalitarian).

The Behavioral System

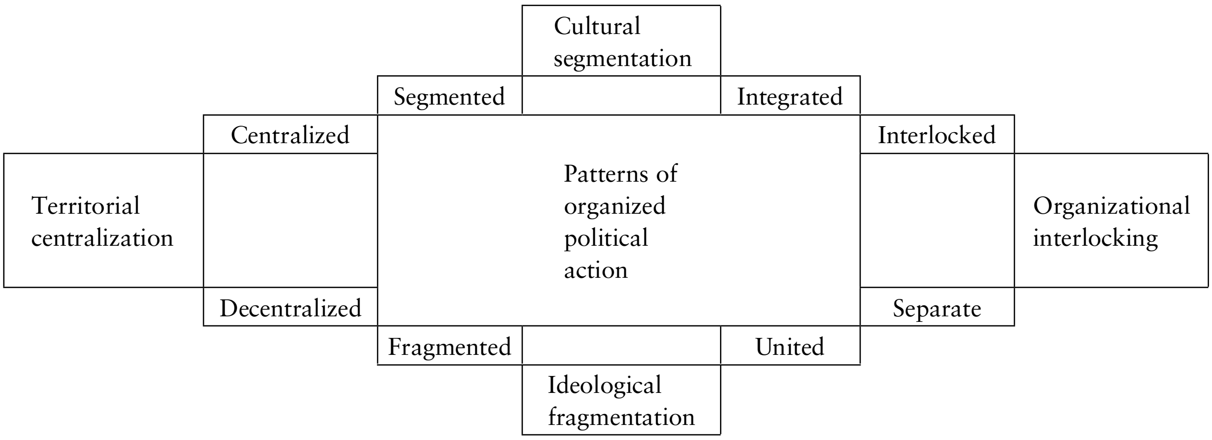
Collective action is undertaken by social movements, interest groups, and politico-electoral organizations or political parties. The levels of rigidity/permeability of the social stratification system and of the normative closure of the identity code affect the predominant form of political behavior and the structure of political agencies.

There is no one-to-one relationship between codes of closure and types of organizational membership. However, some predominant associations can be specified on the basis of the synthetic scheme of organizational membership variations presented in Table 2. Forms of organized political action can be ranged along four dimensions. First, they can be characterized by different degrees of territorial centralization/decentralization (national vs. regional or local forms). Second, forms of organized political action can be analyzed in terms of their degree of "segmentation" along cultural dividing lines (as, e.g., in the case of religious and ethno-linguistic divisions). Third, they can be organized by level of "fragmentation" along ideological dividing lines (as, e.g.,, the political Left–Right fragmentation). Finally, different forms of organized political action can be studied in terms of the level of their "organizational interlocking" with other forms—for instance, the more or less interlocking relationships between political parties, unions, interest groups, or sociopolitical movements.

The presence of primordial codes of closure is likely to generate territorially decentralized or culturally fragmented political organizations and social movements within a country. Civility codes in their pure form are more likely to be associated with the ideological fragmentation of political action, where divisions concern primarily the level, type, and quality of resources distributed within the collectivity and the access to public goods. However, they can occasionally be associated with decentralized forms of political action resting on the specific civic traditions of certain territories. Cultural codes are more encompassing and are conducive to both cultural segmentation (e.g., religious mobilization) and/or ideological fragmentation (progress vs. conservatism, change, and tradition). Well-integrated territorial polities with no internal primordial groups and with high civility and cultural homogeneity tend to produce political organizations that more closely mirror the ideological differences between social and economic interest–based groups.

Therefore, membership- and action-based divisions may be organizationally expressed by territorial decentralization, cultural segmentation,

Table 2 A Map of Political Structuring Variations in Organizational Membership



Source: Adapted from Bartolini, S. (2005). *Restructuring Europe* (p. 103). Oxford, UK: Oxford University Press.

and ideological fragmentation and by different degrees of interlocking relations between specific sociopolitical movements, corporate, and electoral organizations.

Types of Divisions and Cleavages

At this stage, a general typology of all possible divisions and the place that cleavages occupy in such a typology can be achieved by combining the elements of the separate dimensions of the social stratification, normative, and behavioral systems. This is presented in Table 3, which identifies the three bases for all sorts of conflicts and oppositions. They correspond to the interest orientation linked to the social stratification system, the cultural/ideological orientation linked to the normative system, and the organizational base linked to the behavioral system. Taken separately, these oppositions and conflicts identify *simple* divisions. They generate *compound* divisions when they interact and cumulate. In line with the definition at the beginning of this entry, *cleavages* are identified as a special type of complex division that combines these three bases and roots.

Simple Divisions

Some divisions are based simply on processes of interest differentiation and perceptions of interest similarity. Though these interest divisions are mainly categorized as statistical or analytical, they may also occasionally generate conflicts and oppositions. Cultural divides exist among social identities linked to various types of civility and cultural codes (primordial codes are by definition rooted in the social structure via the stigmata of the group), and they have proved their historical resilience even in situations unsuitable for their behavioral expression (e.g., repression). Finally, we can observe conflicts and oppositions that are mainly related to the behavioral divisions among various types of voluntary organizations and that are not linked to divisions rooted in either the social structure or in the normative cultural system.

However, none of these "simple divides" is likely to provide long-lasting sources and bases for the political structuring of the polity unless they are combined. Interest divisions are unlikely to overcome the free riding problem and the cost of organizational mobilization without normative or organizational foundations. Interest orientations need to be "generalized" and somehow made "universal" through an idealized appeal to solidarity. Similarly, pure membership divisions are unlikely to maintain the organizational cohesion of the membership groups only by distributing material incentives and resources. Even the most scandalous and corrupt "spoils" party or organization is likely to require some cultural and/or ideological underpinning to

Table 3 Types of Divisions

		Simple Divides				*Compound Divides*		
Social stratification system	Interest orientation	+	–	–	+	+	–	+
Normative system	Cultural/ideological orientation	–	+	–	–	+	+	+
Behavioral system	Organizational base	–	–	+	+	–	+	+
Resulting divide		Interest divides	Cultural divides	Membership divides	Corporate divides	Social divides	Political divides	Cleavages

Source: Adapted from Bartolini, S. (2005). La formation des clivages [The formation of cleavages]. *Revue Internationale de Politique Comparée, 12,* 32.

survive and stabilize over time. Cultural divisions may be the only ones that are able to survive "underground" with weak interaction bases and loose organizational support.

Compound Divisions

The most important divisions are, therefore, "compound," and they are the only ones able to generate the forms of organized political action we usually observe as social and political movements, interest organizations, and political parties.

The combination of interest orientations with organizational membership may be sufficient to generate highly contingent and goal-oriented organizational cooperation without necessarily being rooted in strong cultural orientations. Indeed, this type of division may allow a considerable amount of within-group differentiation in many other domains different from the one in which the organizational cooperation is generated. This is the typical and elementary form of interest (or pressure) group formation.

Social divisions are based on the combination of interest orientations rooted in the social structure with stronger normative orientations concerning the identification of the group and its role in society. In this type of division, the social determination is less strong than in corporate divisions. The distinctiveness of social roots may be more or less fundamental but the normative orientation is usually essential and decisive. In this case, cooperation is not based on specific and well-defined interests, and the organizational capacity is weak, intermittent, not consolidated, or sometimes nonexistent. The normative orientation is therefore essential to the definition of the group identity. This seems to be the typical form for the emergence of social movements.

The ideal type of political division is expressed via a combination of cultural/ideological orientations with a strong organizational dimension. The ideational solidarity takes the form of common but general goals to be achieved in the polity through the organizational–ideological competition for the attainment of binding decisions. These divisions are more institutionalized and strongly embedded in the institutional structure of the polity. The organizational dimension gives them solidity and stability over time, which may enable them to survive even when the original cultural and ideological bases fade way. Political parties are a typical expression of such political divisions.

Cleavages

Finally, "cleavages" are a type of division that incorporates the three elements of a clear interest orientation, a clear-cut normative vision, and a strong organizational basis. Cleavages encompass corporate and social divisions with political ones. Cleavages show a remarkable degree of resilience over time. They are not transient and intermittent but durable, entrenched, and entrapped. Given their high degree of organizational infrastructure and institutionalization, they often endure over and beyond the interests and the normative orientations that historically gave birth to them. Cleavage lines manifest themselves through an intense interlocking relationship between corporate interest organizations, social movements, and specific electoral organizations. Actually, the distinctive characteristic of cleavages as opposed to all other divisions is that they are characterized by a dense and profound overlapping of membership and exchange of personnel and of resources among social movements, corporate organizations, and electoral organizations. From the organizational point of view, the specificity of cleavage structures is indeed their vertical integration of different forms of organizational action.

As bases of a cleavage, territorial oppositions are more solidly entrenched than cultural segmentations, and the latter are more solidly entrenched than ideological fragmentations. As argued previously, territorial oppositions are rooted in either primordial or civility codes of exclusion, and these have proved to be extraordinarily resilient and impermeable. Cultural codes are by far more open and permeable and therefore both cultural segmentations and ideological fragmentations rest on roots that are historically more malleable and changeable. For example, compared with their impact on ethno-linguistic community definitions, it is more likely that socioeconomic and technological changes make social class positions more permeable and denominational affiliations less significant.

Stefano Bartolini
European University Institute
San Domenico di Fiesole, Italy

See also Class, Social; Ethnicity; Identity, Social and Political; Interest Groups; Parties

Further Readings

Allardt, E. (1968). Past and emerging political cleavages. In O. Stammer (Ed.), *Party systems, party organization and the politics of the new masses* (pp. 66–74). Berlin, Germany: Institut fuer Politische Wissenschaft, Freie Universitaet.

Bartolini, S., & Mair, P. (1990). *Identity, competition, and electoral availability: The stabilization of the European electorate, 1885–1985*. Cambridge, UK: Cambridge University Press.

Eisenstadt, S. N., & Giesen, B. (1995). The construction of collective identity. *Archives Européennes de Sociologie, 36,* 72–104.

Lipset, S. M., & Rokkan, S. (Eds.). (1967). *Party systems and voter alignments*. New York: Free Press.

Rae, D. W., & Taylor, M. (1970). *The analysis of political cleavages.* New Haven, CT: Yale University Press.

Rokkan, S. (1970). Nation building, cleavage formation and the structuring of mass politics. In S. Rokkan (Ed.), *Citizens, elections, parties: Approaches to comparative study of the processes of development* (pp. 72–144). Oslo, Norway: Universitetsforlaget.

Rokkan, S. (Ed.). (with Kuhnle, S., & Urwin, D.). (1999). *State formation, nation building, and mass politics in Europe: The theory of Stein Rokkan*. Oxford, UK: Oxford University Press.

Zuckermann, A. (1975). Political cleavages: A conceptual and theoretical analysis. *British Journal of Political Science, 5,* 231–248.

Zuckermann, A. (1982). New approaches to political cleavages. *Comparative Political Studies, 15,* 131–144.

CLIENTELISM

Clientelism involves a particularistic exchange between patrons and clients in which resources of some kind are traded in a mutually beneficial way. Both parties are supposed to benefit from the exchange, but it involves actors of unequal status. The patrons, for example, landlords or local notables, are the more powerful actors in the exchange. They possess privileged access to resources, such as land, which is denied to their clients, typically poor peasants, migrants, and other marginalized groups. The patrons release these resources in a selective way that directly benefits the clients, and in return, they expect their help, support, or allegiance. These exchanges eventually produce an enduring social relationship, typical for traditional societies, governed by a sense of obligation and loyalty that binds the client to the patron. In modern political settings, clientelism is associated with the selective use of public resources in the electoral process. The patrons, who can be individual politicians or political parties, selectively release various resources—jobs in state administration, titles, pension schemes, or building projects—to individuals or groups in order to "purchase" their vote. This is an understanding dominant in the political science literature where clientelism is seen as a type of representation or of political mobilization that differs from both universalistic programmatic appeals and appeals based purely on personality traits and charisma. This entry focuses on this understanding of clientelism. The first part discusses the concept of clientelism, in particular its relation to other closely related concepts, such as patronage, corruption, and pork barrel. The second part discusses the relevance of clientelism in contemporary political process.

Identifying Clientelism

The passage from traditional to fully mobilized mass democratic societies brought about substantial changes in patron–client relationships. The traditional clientelistic linkage, characterized by face-to-face contact that is managed by a powerful person, has been replaced by exchanges in which an organization, the political party, has become the key distributor of goods and services to the clients. The individual patron–client ties of local notables and powerful landlord has thus eventually evolved into the clientelism managed by complex political organizations. However, studies on the subject have long suffered from a high degree of conceptual vagueness and ambiguity. This problem is particularly evident in the use of the concepts of patronage and clientelism, which are seen either as the same phenomenon or, in contrast, as two different phenomena but without any explanation of what differentiates the two.

Patronage and clientelism do fundamentally denote the same type of exchange but, as Simona Piattoni (2001) points out in the introduction to her edited volume, the latter is a far more penetrating phenomenon. The difference between the two lies in the objects of exchange (i.e., the type of resources that are used in the trade) and in the goals of the exchange (i.e., what the deal is supposed to achieve). Regarding the object of exchange, the key feature of patronage is the discretional allocation of public jobs by (party) politicians. The recipients of these jobs are normally assumed to be party leaders, party activists, influential party backers, or party voters, though the fact that politicians have power for discretional hiring is more important for defining patronage than the characteristics of appointees. In other words, patronage should not necessarily imply that appointees are exclusively party members or party supporters, though in reality it has always been associated with rewarding individuals or groups who have played an important role in party life and party strategy. Concerning the goal of the exchange, parties or politicians are typically seen to allocate jobs in order to secure supporters' votes. However, given that most modern polities are best characterized as having "too few jobs and too many votes," the use of patronage as a way of directly securing electoral support is unlikely to be a very feasible strategy. Instead, patronage has always served mainly organizational and governmental goals, such as entrenching party or politicians' networks within the public sector, serving as an inducement for building and maintaining party organizations by offering selective incentives to party activists in exchange for their work for the party, or ensuring the effectiveness of (party) government by controlling the process of policy making and policy implementation through the appointment of like-minded individuals to key public positions. This understanding of *patronage*—as a distribution of public jobs—coincides with the Anglo-Saxon use of the term. In the American literature, for example, it is widely used to describe the politics of (city) machines in which bosses handed out patronage jobs to shore up their campaign organizations; it is also employed to describe the wholesale changes in administration following presidential elections, sometimes referred to as the "spoils system."

In contrast to patronage, clientelism should essentially be considered an electoral tool in which benefits are delivered to obtain the recipients' vote. In fully mobilized mass societies, clientelism denotes an exchange between a political party on the one hand and individuals or groups on the other hand, in which divisible benefits are offered directly and on a wide scale to tie supporters to the party. The object of these exchanges can of course be public jobs, but this is unlikely to secure any major electoral advantage to patrons, as argued above. Clientelism most often includes distribution of a wide variety of other benefits, ranging from food and medicines to a pension or a low-interest loan, access to housing, or award of selective development projects in a particular area. Therefore, in contrast to patronage, clientelism, especially in its mass society incarnation, requires other methods of reaching the clients than just a distribution of a handful of state jobs. These methods are often fairly bureaucratized and based on impersonal measures, such as when the incumbents pass a decree favoring a particular category of persons, such as taxi drivers, teachers, or small shopkeepers. Clientelism also implies an asymmetrical nature of the linkage, which takes place between actors of different status and power. Even when both sides accrue benefits and both may perceive the exchange as mutually beneficial, the clientelistic linkage often entails an element of inequality. This is also why, empirically, clientelism is most often found in the context of widespread urban and rural poverty and inequality. In contrast, distribution of patronage jobs to party activists or key party benefactors is likely to involve a far more symmetrical relationship of power, at least in that the recipients will be less dependent on obtaining these jobs, and their social status will often be equal to that of their political patrons.

Pork barrel politics is normally subsumed as a subtype of either clientelism or patronage. Yet it is a distinct practice that denotes tactical allocation of government funds, usually in the form of legislation on public works projects, to favor specific constituencies. Susan Stokes (2007) distinguishes pork barrel from clientelism on the basis of their different distributive criterion. While the distributive criterion of clientelism is "Did you (will you) vote for me?", the implicit criterion in the distribution of pork barrel politics is "Do you live in my district?" In other words, while clientelism entails a benefit for particular individuals or groups, pork barrel implies that a whole constituency is favored by a public policy decision. Although the goal of both clientelism and pork barrel politics is to obtain the recipients' electoral support, they also differ in that the element of exploitation and inequality that characterizes the former is absent in the latter. That is probably the reason why these two practices are viewed differently in normative terms. Politicians who deliver goods and services on a clientelistic basis usually try to keep it as a secret matter between them and the clients. In contrast, politicians who manage to pass pork barrel legislation are often eager to present them as a political asset.

Indeed, both patronage and clientelism are nearly always seen as corrupt or morally questionable practices, even though corruption is empirically and conceptually a distinct phenomenon. Political corruption denotes illegal public decisions taken by parties or individual politicians to obtain financial resources, such as when parties may favor firms by handing over the concession of a public utility or permitting the development of an economic activity without the fulfillment of all legal requirements and demand, in exchange, financial kickbacks to party coffers. Clientelism and patronage can undoubtedly lead into corruption, such as when discretionally appointed party hacks use the state office for personal enrichment through demanding bribes or when politicians try to obtain resources for their clientelistic policies through illegal dealings with economic firms. In that sense, both patronage and clientelism can breed corruption, which might explain why the public discourse in many postcommunist countries refers to clientelism in connection with the penetration of the state by organized crime and by other rent-seeking groups. However, patronage appointments are often legally sanctioned and above board, and hence not corrupt; similarly, many of the clientelistic deals between patrons and their voters and supporters fit, even if uneasily, within the legal remit of policy and administrative decision making. Importantly, clientelistic exchanges are never directly aimed at obtaining monetary goods on the part of the patrons.

To recap, clientelism and patronage, though closely related, can usefully be distinguished on the basis of the goals and objects of the exchange. Clientelism—a particularistic form of political and electoral mobilization—is by definition more penetrating than party patronage, usually reaching larger numbers of people and covering wider ranges of exchanges. However, patronage—discretional allocation of jobs—is the necessary (but not sufficient) condition for the emergence of clientelism, for it is mainly due to their ability to control public jobs that parties and politicians are able to manipulate administrative processes in order to deliver targeted divisible benefits to their clients. In other words, insofar as parties or politicians do not control at least the key jobs in the public administration, they will hardly be in the position to develop large-scale "clientelistic" exchanges. This is why patronage and clientelism will always remain closely related in empirical reality, in spite of the fact that handing out state jobs will lead to tangible electoral returns only in very small democracies or in countries with limited suffrage and partial mobilization.

Clientelism in Contemporary Democracies

Traditional clientelistic practices have usually been seen as a product of early-modern development and to be particularly important in rural and less economically well-developed regions and in political systems in developing countries. For numerous reasons, they were expected to disappear in the process of political and economic modernization and development of mass party organizations that characterized the 20th century. First, the individualized processes of vertical integration and mobilization associated with patron–client ties were undermined by the forms of horizontal mobilization and integration promoted by class politics and cleavage structuring. In other words, as a society developed mass electoral alignments, the scope for individually based networks of supporters became more limited. Second, as the mass party itself took over from premodern and cadre parties, and as appeals based on programs and ideology replaced those based on more personalized political representation, it also became more difficult to sustain patron–client links, and especially to build a distinctive clientele. Third, as polities modernized and professionalized, meritocratic systems of advancement became more acceptable and widespread, and hence objective rules, exams, and qualifications replaced favors, friendships, and networks in the process of career building. Here too, then, the scope for patronage as one facet of clientelism became limited, in this case as a result of a shortage of supply. Fourth, as economies modernized, local markets, especially in poor rural areas, became better integrated into larger regional, national, or even supranational units. The members of these initially isolated communities entered into commercial and social ties with persons in other places, which, in turn, diminished demand for patrons to advance their interests at the center.

Finally, as societies also became richer and more educated, the demand for particularized benefits or favors diminished, while citizens also became more confident of their own ability to deal with the bureaucracy. In this sense, citizens began to handle their own affairs more effectively and hence have less need of a patron to work on their behalf. As a result, the deals between patrons and clients also became increasingly difficult to monitor and enforce. Now it was no longer possible to scrutinize clients' behavior through their natural seclusion in relatively small and homogeneous communities; in addition, the increased mobility and education among the electorates made the enforcement of clientelistic deals nearly impossible.

All of this implies a more or less secular process, whereby clientelism becomes steadily eliminated in favor of collective representation and accountability. The key actors in this process are, of course, the emerging mass parties, which encourage horizontal rather than vertical integration and which promote the provision of universal rather than particularistic benefits. However, this is not to suggest that clientelism did not survive the advent of the mass party. In many countries, for example, Italy, Austria, Ireland, and Japan, mass parties have also been known to adapt these practices as a means of ensuring their own electoral survival. Rather than making a full leap from traditional clientelistic to programmatic politics, as was widely expected in the 1950s and 1960s, the mass parties in these countries instituted a relatively bureaucratized and impersonalized system of clientelism, which combined particularistic politics with the more universalistic appeals. This was no longer the personalistic patronage system based on organic ties between patrons and clients but still a system in which a large proportion of public services and jobs was distributed to selective clienteles in an attempt to build political support and win elections.

What is clear, however, is that clientelistic practices in general have gradually become more difficult to sustain in the era of the mass party and hence proved more exceptional than conventional in modern democracies. Indeed, bureaucratic clientelism became difficult to maintain even in the advanced democracies in which it functioned well into the 1970s and 1980s. This happened mainly for two reasons. First, parties and politicians today face more and more constraints on acquiring clientelistic resources. With the shrinking of the public sector due to privatization and marketization on the one hand, and increasing exposure of domestic economies to pressures of globalized markets on the other hand, it is now more difficult for politicians to obtain resources for greasing the clientelistic system without compromising performance of the economy and inflow of investments. Second, the shrinking of resources is also related to the vigilance of mass media and civil society groups in reporting clientelistic and other abusive state practices, on the one hand, and the increased willingness of judicial authorities to punish them, on the other. These may not be universally effective in all countries, but it is clear that anticorruption programs together with various measures to limit the politicization of the public sector have constrained the range of options that clientelistic parties and politicians can exploit at their discretion. All in all, therefore, whether one looks at the supply or demand side, in contemporary modern polities, the tide seems to operate against clientelism.

Does this mean that clientelism is irrelevant in contemporary societies? If one looks at the upsurge of works on clientelism in the recent decade, such as Piattoni (2001) and Herbert Kitschelt and Steven Wilkinson (2007), the answer must be a resounding no. These new studies suggest that far from being dead, clientelism is likely to be an enduring feature of politics and to encourage inquiries into the conditions that favor or constrain clientelism in contemporary democracies. This is in no small part thanks to the recent waves of democratization in Latin America, South Asia, Eastern Europe, and Africa. In many of these emerging new democracies, clientelistic practices were well-entrenched phenomena prior to democratization and hence were always likely to play a role in the quest for power and votes after the (re)entry of competitive politics. These newly democratized polities are often marked by the emergence of weakly structured parties and party systems that are unlikely to attract voters (only) on the basis of coherent programmatic packages. Recent developments in the patterns and processes of politics in advanced industrial democracies also suggest at least a partial rethinking of the familiar assumption about the relative insignificance of clientelistic practices as a mode of representation. It can, for example, be argued that the declining intensity of ideological differences across parties in many modern democracies may encourage cultivation of clientelistic links at the grassroots level at the same time as it encourages a more personalized style of leadership at the top.

In other words, whether one looks at contemporary new or long-established democracies, it seems that political clientelism has received a fresh impetus and hence might prove to be, somewhat paradoxically, the product of modernization rather than its victim. Few case studies aside, however, there is still a lack of solid cross-national empirical evidence that would probe into these theoretical expectations concerning the resurgence of clientelism and show how important it is as a mode of citizen–politician linkage. The evidence is clearer with respect to patronage, which was distinguished here from clientelism on the basis of the type of resources that are traded and the goals that it serves to politicians. Like clientelism, patronage politics can also prove resilient and adaptable. Indeed, patronage is still likely to prove an effective strategy for dealing with problems of party organization and party building, particularly as the traditional representational links between parties and society become weaker. In other words, as modern parties become more entwined within the institutions of the state, and as they lose their traditional grounding within the wider society, patronage can become a key resource in anchoring the party presence within the political system and in controlling flows of communication. Through the appointment of party personnel to key agencies and institutions, parties can hope to gain an oversight of the likely demands posed to political leaders as well as of the likely policies and programs that are needed to meet these demands. Patronage can therefore compensate for otherwise decaying organizational networks.

Several studies focusing on both new (e.g., Anna Grzymala-Busse, 2007; Conor O'Dwyer, 2006) and established democracies (e.g., Guy Peters & Jon Pierre, 2001) have effectively shown that political use of public jobs is a rampant practice in many of these polities; these studies also ask critical questions about factors that might explain cross-national differences in the pervasiveness of patronage practices. However, it is probably fair to say that political science literature has dealt with the development of patronage networks within the state usually only in passing, under the banner of party government (see, e.g., Jean Blondel & Maurizio Cotta, 1996). Moreover, as Dominic Bearfield (2009) correctly points out, the negative connotation associated with patronage has led academics to understand patronage as corrupt and pathological, and hence something either undesirable or difficult to study empirically. It is for these reasons that many excellent studies dealing with patronage appointments can be found in public administration rather than in political science literature, even though the specific concern of that literature is usually less with the role of parties and politicians in the process of patronage appointments and more with the impact that it has on the performance of bureaucratic organizations. Largely as a result of this disjuncture between public administration and political science analytical objectives, here too we still lack solid cross-national evidence that would allow us to scrutinize the range and depth of patronage practices and lead us to appreciate the precise goals that political patrons seek to achieve in making patronage appointments. This represents a significant gap in the literature on the workings of democracy, political parties, government, and institutions of the state because theoretical reasons for the survival and adaptation of patronage practices in contemporary democracies are perhaps even more compelling than the expectations concerning the relevance of clientelism.

Petr Kopecký
Leiden University
Leiden, Netherlands

See also Mobilization, Political; Party Linkage; Party Organization; Politicization of Bureaucracy

Further Readings

Bearfield, D. A. (2009). What is patronage? A critical reexamination. *Public Administration Review, 69,* 64–76.

Blondel, J., & Cotta, M. (Eds.). (1996). *Party and government: An inquiry into the relationship between governments and supporting parties in liberal democracies*. Houndmills, UK: Macmillan.

Eisenstadt, S. N., & Lemarchand, R. (Eds.). (1981). *Political clientelism, patronage and development*. London: Sage.

Grzymala-Busse, A. (2007). *Rebuilding leviathan: Party competition and state exploitation in post-communist democracies*. Cambridge, UK: Cambridge University Press.

Kitschelt, H., & Wilkinson, S. (Eds.). (2007). *Patrons, clients, and policies: Patterns of democratic accountability and political competition*. Cambridge, UK: Cambridge University Press.

O'Dwyer, C. (2006). *Runaway state building: Patronage politics and democratic development*. Baltimore: Johns Hopkins University Press.

Peters, G., & Pierre, J. (Eds.). (2001). *Politicians, bureaucrats and administrative reform*. London: Routledge.

Piattoni, S. (Ed.). (2001). *Clientelism, interests, and democratic representation: The European experience in historical and comparative perspective*. Cambridge, UK: Cambridge University Press.

Shefter, M. (1994). *Political parties and the state: The American historical experience*. Princeton, NJ: Princeton University Press.

Stokes, S. (2007). Political clientelism. In C. Boix & S. Stokes (Eds.), *Handbook of comparative politics* (pp. 604–627). Oxford, UK: Oxford University Press.

Coalitions

A coalition is the temporary cooperation of different individuals, groups, or political parties to achieve a common purpose, which can be either short term or long term. Almost all politics can be conceived as involving the formation of some kind of coalition. Pressure groups, social movements, and political parties are coalitions of individuals with a common interest; governments can be formed by coalitions of political parties, not only in parliamentary regimes but also in regimes of division of powers; maintaining political stability or resolving political or ethnic conflicts can induce the formation of broad coalitions committed to support a new regime; and different governments and states can form military coalitions unified under a single command. A coalition implies cooperation among its members. A relationship of conflict usually develops between different coalitions whose members have opposite interests.

The most usual analyses of coalitions in politics deal with the formation of multiparty cabinets in parliamentary regimes. In government coalitions, several political parties cooperate, usually during a legislative term between two elections. Coalition governments are the most usual form in most parliamentary regimes using electoral rules of proportional representation, which typically do not produce a single-party majority of seats. This includes most countries in continental Europe, as well as other democratic countries such as India, Israel, and New Zealand. However, other broad coalitions have resulted from political settlements or conflict resolution, such as in South Africa, Zimbabwe, Kenya, or Madagascar, often involving international or external moderators.

An example of government formation in parliamentary regimes is Germany, where all cabinets since the end of World War II have been multiparty coalitions. The chancellor or prime minister has always been either a Christian Democrat or a Social Democrat, but government coalitions include several formulas: rightist Christian Democrats (always with their allies, the Bavarian Social Christians) with center-right Free Democrats, center-left Social Democrats with Free Democrats, Social Democrats with left Greens, and the so-called "grand coalition" of Christian Democrats and Social Democrats. All these governmental formulas have had majority support in terms of both popular votes and parliamentary seats. Due to the long-term participation of a few parties in government and other institutional mechanisms, the degree of stability in major public policies in Germany is very high. In contrast, the alternation of single-party cabinets when different parties receive a majority of seats in successive elections, as in Britain, may provoke periodic shifts in major public policy making.

Within political science, the formation and termination of cabinet coalitions have been analyzed as cooperative games in game theory. Both the search for office and the search for policy or ideological goals can be presumed to be realistic and legitimate motivations of politicians when they try to form a multiparty coalition. At the time of forming government, the interest of members of parliament in enjoying as much power as possible translates into the aspiration to accumulate as many government portfolios or ministries as possible for their party. This becomes a criterion to form a coalition with the minimum viable size. The explanation for this is that if a government is formed of a multiparty coalition without superfluous members, it can give each party a relatively high share of power to exert and enjoy. In a minimum winning coalition, each of its party members is pivotal, in the sense that the loss of a party would render the coalition no longer a winning coalition. In the particular case when a party has an absolute majority of seats in parliament, the minimum winning "coalition" is the majority single-party government without additional partners.

For parties interested in policy, a criterion to select potential partners in a government coalition is the minimization of policy–ideology distance. Specifically, parties may try to form a coalition with "connected" parties, for example, on the left–right dimension, that is, with parties that are contiguous to their positions and then devoid of unnecessary parties. Closeness can facilitate the negotiation of a government program and diminish internal policy conflicts within the coalition. For instance, social-democratic, leftist, and green parties are more likely to form coalitions among themselves than with liberal, Christian Democratic, or conservative parties, which in turn can be prone to unite themselves in some governmental coalition.

A minimal connected winning coalition with more than two parties may include some superfluous partners in terms of size that are located on intermediate ideological positions and are thus necessary to maintain the ideological connection between its members. But also in a minimal connected winning coalition, each party is pivotal because the loss of a party would render the coalition either no longer winning or no longer connected. On a single-dimension policy space such as the left–right axis, the median party will always be included in a connected winning coalition. The median is the position having less than half positions on each side and is, thus, necessary to form a consistent majority along the issue space. Empirical analyses show that government coalitions containing the median party in parliament are more likely to be formed.

When different winning coalitions can be formed in a parliament, the party composition of the government may depend on the bargaining power of each party and the presence, or not, of a dominant party. Several tools can be used to analyze these points. Different political parties may have bargaining power to form a government coalition, which does not mechanically correspond to their numbers of seats. A relatively minor party, which is "pivotal" to form a majority, that is, a party able to contribute with the necessary number of seats to make a coalition winning, may have relatively high power to negotiate cabinet membership or policy decisions. For example, some center, agrarian, ethnic, radical, or democratic parties may be located in a central place able to form coalitions, at different moments, with either the parties on their left or those on their right. In contrast, a relatively large party whose contribution can be easily replaced with that of a smaller party may have relatively low bargaining power in comparison to its size. A party's bargaining power in parliament can be measured, thus, not by its number of seats but by the proportion of potential winning coalitions in which the party is pivotal. There are several "power indices" available to measure a party's bargaining power. They slightly differ in their assumptions regarding actors' criteria, coalition models, and decision rules, but most of them produce similar results.

In certain configurations, the largest party in parliament, even if it does not have a majority of seats, can be dominant if its central position makes it able to block any coalition cabinet and take all portfolios. In other words, a party is dominant if the other parties cannot form a winning coalition without that party. If the nondominant parties are unable to form a majority coalition among themselves due to their ideological distance, then the dominant party can block any coalition and form a minority cabinet. Minority cabinets of a dominant party are viable and likely to be formed the greater the policy–ideology divisions and the smaller the size of the parties in the opposition. For example, the Indian National Congress party in India, the Christian Democratic Party in Italy, the Liberal Democratic Party in Japan, and the Social Democratic parties in Denmark, Sweden, and Norway were dominant single parties for long periods.

Additional analyses deal not only with which parties are more likely to enter a coalition cabinet but also with the allocation of ministries to the parties within the cabinet. The distribution of cabinet portfolios among coalition parties tends to be proportional to the number of seats controlled by each party—that is, with its contribution to making a coalition winning. However, different parties have preferences for different portfolios depending on the policy issues they emphasize the most, which may produce varied allocations. The prime minister's party usually controls most of the portfolios in charge of major policy domains, especially economy, defense, and home affairs. Other cabinet portfolios can be allocated to parties with a strong profile on certain issues on which they tend to campaign and attract citizens' votes, such as social policy for Laborites or Social Democrats, education for Christian Democrats, finance for Liberals, agriculture for Agrarians, and so on.

Different coalition formulas are associated with different durations of cabinets. Regular parliamentary elections are usually scheduled at intervals of 3, 4, or 5 years, depending on the country. But a significant number of parliamentary cabinets do not last as long as they legally could because there are anticipated dissolutions of parliament, elections (which can usually be called by the prime minister), resignations by prime ministers, and successful motions of censure and defeated motions of confidence. Regarding the party composition of cabinets, single-party majority cabinets tend to last longer than multiparty coalition or minority cabinets. In single-party governments, conspiracies among party members to replace the incumbent prime minister are relatively likely, especially if party members expect to have better electoral prospects with a new candidate. In contrast, in multiparty coalition governments, internal party cohesion tends to increase, but coalition partners are more willing to work against the incumbent formula. For coalition cabinets, the higher the number of parties and the broader the ideological distance between them, the more vulnerable to splits and departures and less durable they should be expected to be. This kind of crisis is relatively frequent in parliamentary regimes. In 15 countries of Western Europe since World War II, about one sixth of parliamentary governments have not concluded their term due to a change of prime minister, a change in the party composition of the government coalition, or the dissolution of parliament and the call of an early election.

Multiparty coalition cabinets in parliamentary regimes tend to induce a relatively balanced interinstitutional relationship between the prime minister and the parliament. As political parties need to bargain and reach agreements in order to make policy decisions and pass bills, they learn to share power and develop negotiation skills. Cabinet members from different parties need to cooperate as well. The prime minister cannot prevail over the cabinet or the assembly as much as when leading a single-party government because, even if he or she is a member of one of the parties involved, he or she has to negotiate with the other parties and maintain the coalition united. In a parliamentary regime, the institutional role of the parliament thrives when no party has an absolute majority of seats.

Josep M. Colomer
Institute for Economic Analysis, Higher Council for Scientific Research
Barcelona, Spain

See also Cabinets; Cooperation; Game Theory; Parliamentary Systems; Parties; Party Systems

Further Readings

Laver, M., & Schofield, N. (1990). *Multiparty government*. New York: Oxford University Press.
Laver, M., & Shepsle, K. (1997). *Making and breaking governments*. New York: Cambridge University Press.
Lijphart, A. (1999). *Patterns of democracy*. New Haven, CT: Yale University Press.
Muller, W. C., & Strom, K. (Eds.). (2008). *Cabinets and bargaining*. New York: Oxford University Press.
Riker, W. H. (1962). *Theory of political coalitions*. New Haven, CT: Yale University Press.

Cohort Analysis

A typical question posed by researchers of political socialization is whether people socialized under distinct political and social conditions manifest different attitudinal and behavioral patterns and the extent to which such differences persist during their life span. Addressing this question at the individual level is difficult because it requires tracing different age-groups during their life trajectories. Since long-term panel data are still scarce in political science, a common approach is to focus on aggregate patterns of change. The accumulation of cross-sectional data during recent decades has enabled the use of pseudopanels, created by pooling repeated cross-sections over a long period of time. Accordingly, units are aggregated after being classified according to their age and the period in which the phenomenon of interest is measured. For example, if the aim is to examine the sources of variation in turnout levels over the past 50 years, 25-year-olds interviewed in the 1970s can be simultaneously compared with 50-year-olds interviewed in the same year interval and with 25-year-olds interviewed in the 1990s. In so doing, the aim of this type of analysis, known as cohort analysis, is to explore whether variation in an attribute of interest is mainly accounted for by particular political periods, age-groups, or subcategories defined by an interaction of the two. The aim of this entry is to describe this analytical strategy and to shed some light on some of its limitations.

Period Effects

As a way to motivate the discussion, let us imagine a population of interest, voting for one of the two parties, A and B, which compete in elections that are held every 4 years. Assuming that the only source of change in the composition of the population is that of succession, that birth and death rates are approximately equal, and that the structure of the party system remains stable, the question is how can political change occur in this context? It turns out that change regarding a social attribute of interest under these circumstances is essentially produced through two different mechanisms, both of which can be traced in Table 1, which distinguishes only between two subgroups of interest, namely, young and old voters (the reason for using age as a forcing variable will be clear in the next paragraph). The first mechanism involves a coming event or a sequence of events that might induce people to change their minds. This is referred to as a period effect. For instance, people might vote predominantly for Party A during Time 1. An event takes place, and they shift to Party B. If this shift occurs at approximately equal rates between different age-groups, one should observe a difference between the two columns of Table 1, whereas no difference should be found between the two rows of each column. This is shown in the upper panel of Table 1.

Cohort Effects

Political events, however, are unlikely to affect all people in the same way. Most of the time researchers

Table 1 An Illustrative Example of Period, Cohort, and Life Cycle Effects

	Time 1	*Time 2*
Period effects		
Young	A	B
Old	A	B
Cohort effects		
Young	A	B
Old	A	A
Life cycle effects		
Young	B	B
Old	A	A

observe (or assume) particular cohorts, that is, groups of individuals who pass some crucial stage of their lives at approximately the same time, to be considerably more affected by political events than other cohorts. Although different cohorts of interest have been examined, the most common approach is to distinguish between new cohorts and older ones with the first deemed more susceptible to political stimuli. That this assumption is implicitly made in most cohort analyses is evident from the definitions given to different cohorts so as to distinguish one from the other. In most cases, each cohort is characterized by the salient historical and/or political features that marked its early adulthood (the "war generation," the "1968 generation," etc.). If such shaping socialization processes dominate over inertial forces of intergenerational transmission of political attitudes, young adults can differ in terms of their value orientations or any other attitudinal aspect of interest from one generation to another. This second mechanism of change, known as a cohort effect, is presented in the second panel of Table 1. Whereas the older age-group remains stable over time, young cohorts shift from Party A to Party B. According to this scenario, change is driven only through the entrance of new generations.

Life Cycle Effects

Political change can hardly ever be attributed to either of the two sources independently. In most cases, the two mechanisms interplay and the evaluation of their independent contribution in aggregate change is far from a straightforward enterprise. This effort is further blurred by the role of life cycle. The latter can incorporate physiological, social, and psychological intracohort changes taking place during chronological aging. Physiological changes may have important attitudinal implications (e.g., participation in politics becomes more costly for the elderly) but they tend to vary only slightly among different political contexts, and thus, they do not constitute the primary source of interest in most political science applications. Social change due to aging involves contextual influences stemming from peer groups and social roles (getting married, finding a job, settling down, etc.) that most people assume during his or her life span. These influences might, for example, lead young people to opt for a more liberal party than their older counterparts. Such evidence would be the outcome of an age gap, largely persistent through time. This is seen in the final panel of Table 1, which indicates an age difference in party choice that holds at both points in time. This pattern cannot account for change over time unless the population distribution changes in a consistent fashion. For instance, if older people opt for more conservative parties than younger people, a gradual increase of their relative portion in the composition of the population might reveal an aggregate difference over time without people actually changing their attitudinal profiles. Such a pattern would be due to aging effects.

Psychological change can manifest itself in various ways. For instance, disengagement theory asserts that some otherwise distinct life stages might for different reasons discourage people from participating in political affairs. More important, aging might serve as a proxy for the process of reinforcement and crystallization of people's prior orientations. This is probably the most salient effect of the life cycle, especially in applications of voting research, and has been already implicitly alluded to in the previous section. People form their political predispositions in a developmental manner, and thus, change in attitudinal respects is increasingly less likely to occur as people accumulate experience with the political world. The example provided below shows how this process might confound the observed pattern of aggregate change.

An Example

Let us now assume that in our imaginary population we observe a trend shown in Figure 1, that is, the percentage of the population affiliated to a party during the past 30 years decreases in a rather progressive fashion. What could we infer about the sources of change? On the one hand, people's attachments with the parties might have declined as a consequence of disappointing government performance. On the other hand, the coming of new cohorts that are on average less attached to the parties might also account for this pattern. Looking only at Figure 1, it is impossible to say which of the two explanations is at work. Consequently, a typical first step in such types of analysis is to disaggregate this general trend into different cohorts. Figures 2 and 3 do that, each one giving pride of place to one of the two mechanisms described previously.

As can be seen, the two figures tell a completely different story. In Figure 2, pure period effects are observed. Each new cohort enters the electorate from the exact point that the previous cohort stopped and they continue together thereafter. No intercohort difference is observed and hence the cohort-specific curves merge in a single curve, making the graph look identical to Figure 1. In other words, distinguishing between age-groups is redundant because all of them react identically to the political influences they experience. Figure 3,

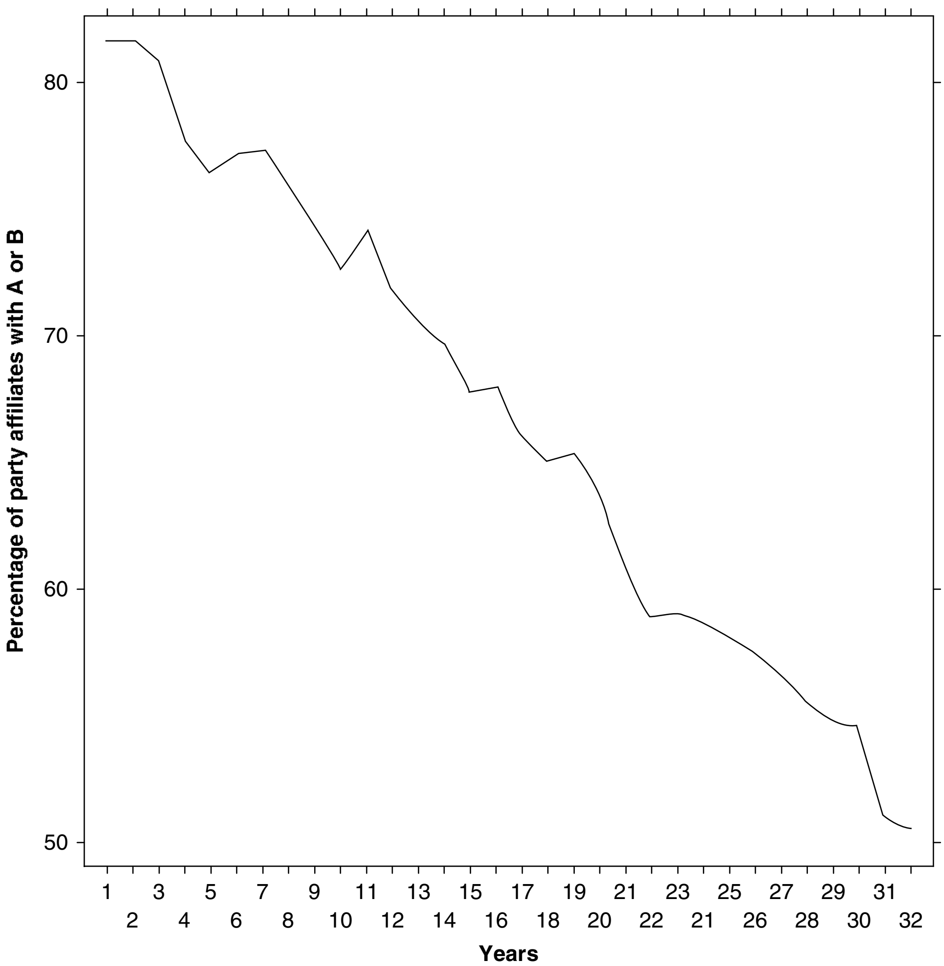


Figure 1 Percentage of People Registering Partisanship With Either Party A or Party B in a 30-Year Time Interval

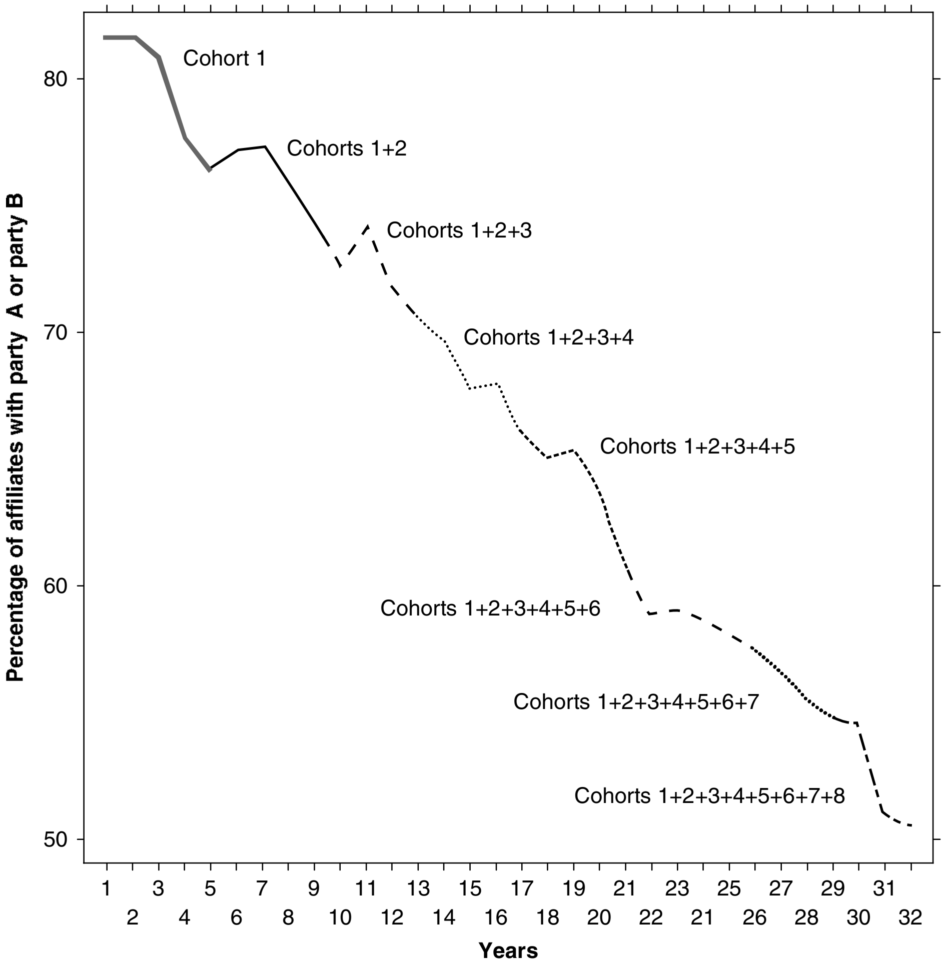


Figure 2 Period Effects

in contrast, shows that all cohorts remain relatively stable through time but each coming cohort is constantly less affiliated to either of the two parties. Consequently, the overall trend is simply the outcome of intercohort variation, although each cohort designates remarkable stability in its political orientations during the observed time interval. Political change, in this case, is simply due to the lower starting level of each coming cohort. If that was what produced the average decline, change would need to be allocated to pure cohort effects.

Needless to say, the artificial lines drawn here only serve as idealized benchmarks useful to evaluate the role of each of the two mechanisms against real data. That said, social reality is much more complicated since both mechanisms, albeit at different rates, are typically at work. To make matters worse, any analysis of such effects needs to account for the role of aging, which can confound either cohort or period effects. In the first case, if aging implies a societal type of change, not taking it into account might give room to spurious cohort effects. Addressing this possibility requires gathering data over a long period of time. This is the only way to examine whether differences between young people and older cohorts persist over time. For instance, to see whether the fact that 20-year-olds in the 1970s vote at lower levels than 40-year-olds is due to aging or cohort effects, one needs to

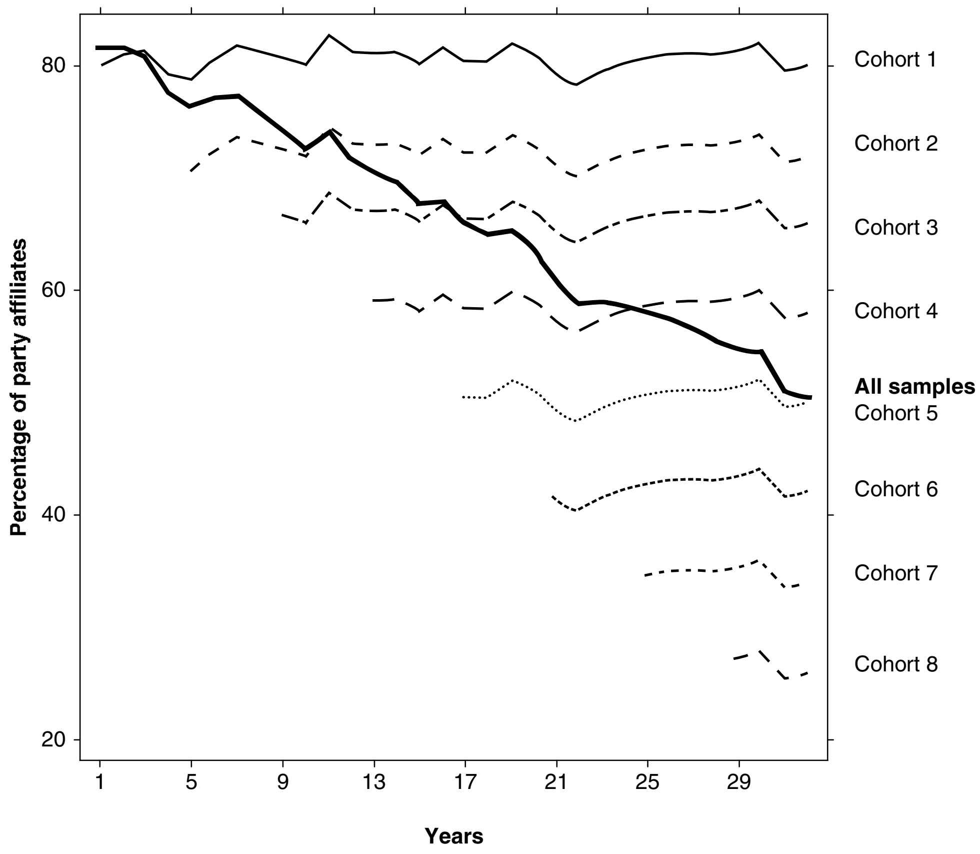


Figure 3 Cohort Effects

Note: There is always some noise (error) in the curves, indicating that even in the most ideal case, some fluctuation would still exist due to sampling variability.

trace both groups for at least 20 years. If both groups continue voting at the same levels as 20 years ago, we would observe evidence in favor of cohort effects. If the first group approaches the turnout levels of the older group, this would constitute evidence for life cycle effects. The problem, of course, is that on many occasions these mechanisms might coexist, making their independent contribution in the overall pattern of change difficult to assess.

Aging can also be a confounder or a mediator of period effects. This is mainly through the immunization hypothesis. To explore whether aging implies an increased resilience to political change stimuli, we need to compare the trajectories of different cohorts over time. This is done in Figures 4 and 5. In the first case, no particular period shock is assumed. The first group, representing people with an already long record in political affairs, seems to remain stable over time. Newly coming cohorts, however, begin from a low starting level but progressively strengthen their feeling of belonging to a particular party as they accumulate electoral experience. Failing to account for this intercohort difference might falsely result in attributing the change to period effects. To be sure, there are ceiling effects in this process that make further differences due to age evaporate. This pattern can be grasped by the comparison of Cohorts 6 and 7 of Figure 4, where initial age-related differences diminish as both cohorts increase their level of attitudinal stability. The only difference between the two cohorts is that the more recent one began its trajectory with a lag. However, their

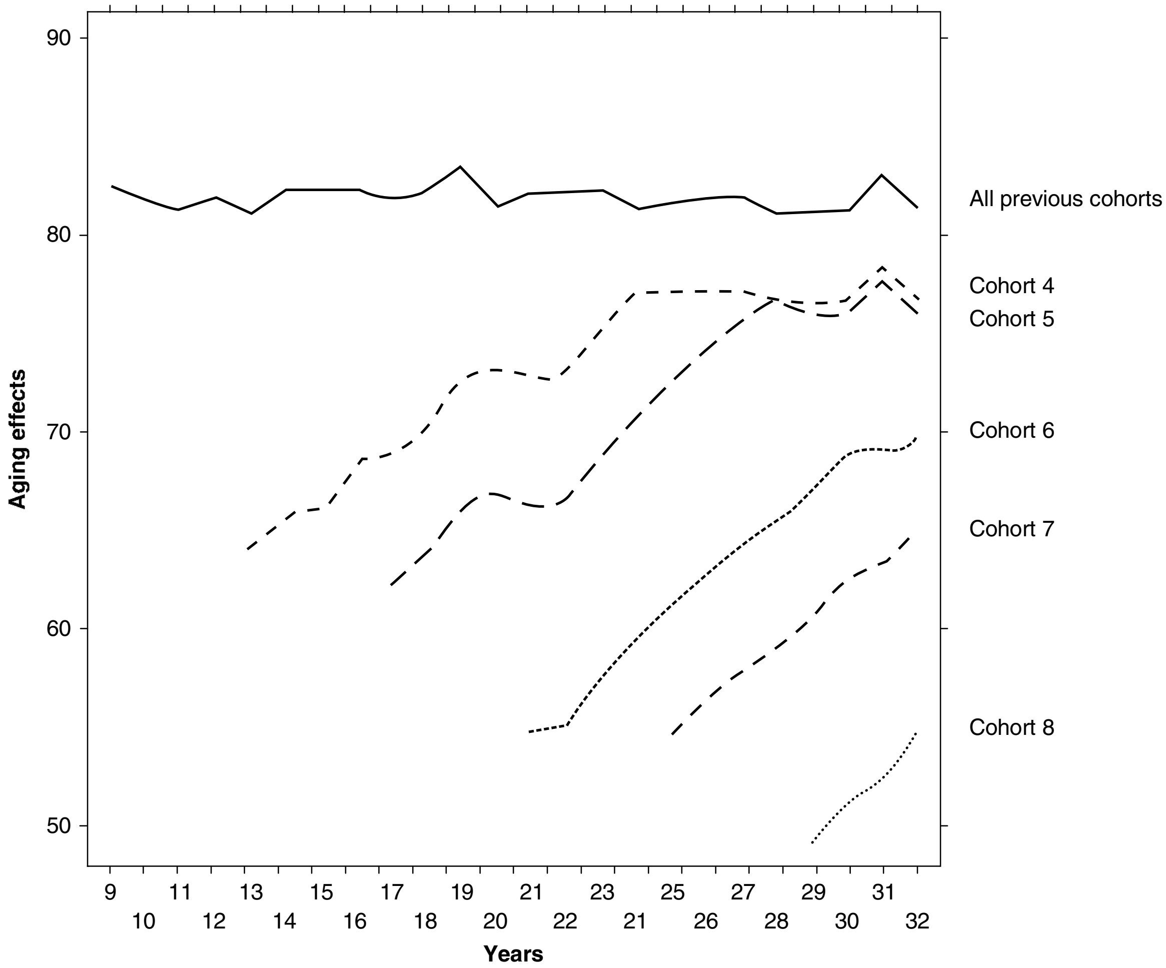


Figure 4 Aging Effects I

route is otherwise indistinguishable. In the presence of ceiling effects, this means that they eventually converge once this process of partisan learning is complete.

Figure 5 shows how aging can work as a mediator of period effects by approximating the idea that period shocks are more consequential among younger cohorts. A temporal shock causing a decline in the level of partisan affiliation in Year 28 is more decisive for people with lower levels of prior electoral experience. As can be seen from the downward spike, the magnitude of this effect diminishes in a monotonic fashion as we move from younger to older cohorts. The reason for that, once again, is that the latter experience this shock having already established firmer political beliefs.

The Age–Cohort–Period Framework and Its Deficiencies

A systematic attempt to simultaneously address all three potentially confounding factors is next to impossible. Until recently, the industry standard in this subfield was to disentangle the sources of political change through the so-called Age–Cohort–Period (ACP) framework. Four important limitations characterize this procedure. The first problem is the well-known identification problem. Since each of the three terms is a linear function of the

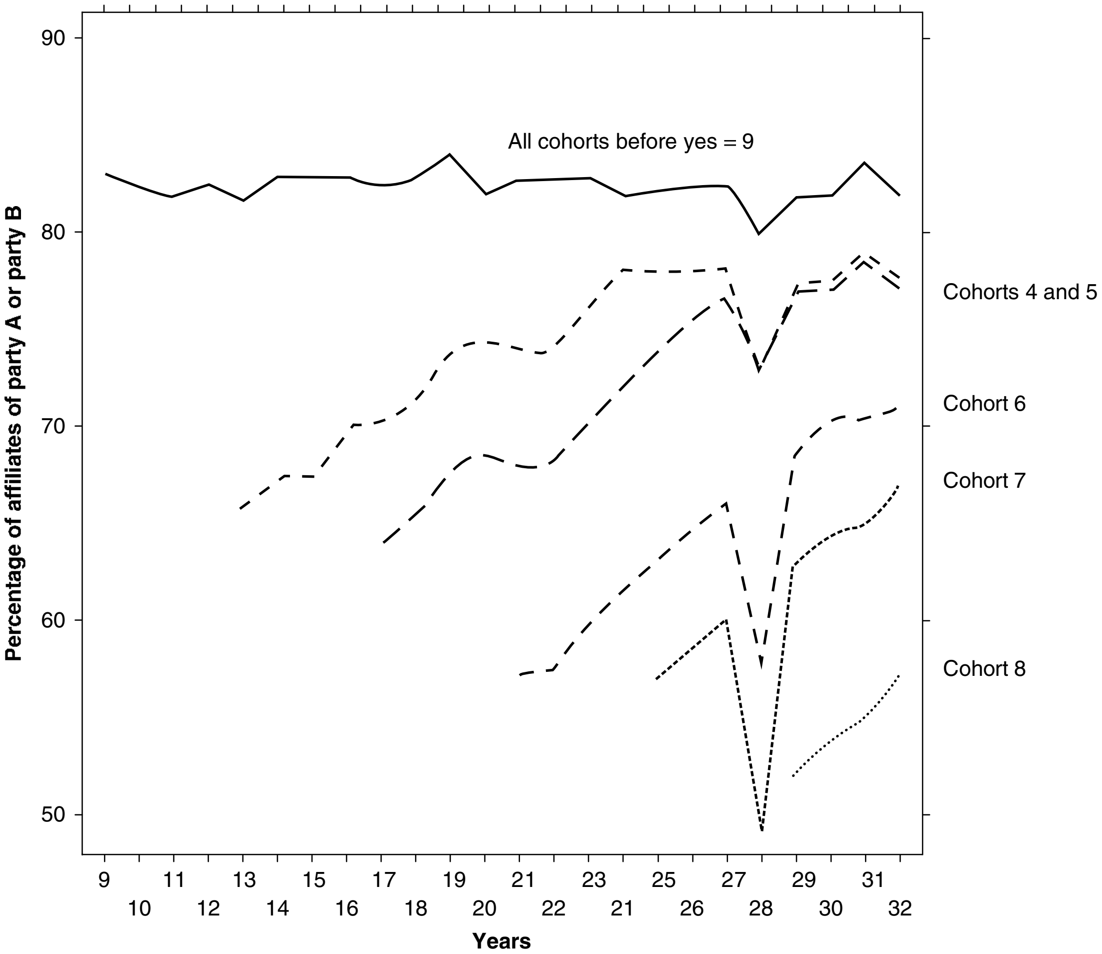


Figure 5 Aging Effects II

Note: To retain simplicity, each group is allowed to return to previous levels immediately after the shock, hence it is assumed that there is no difference in the long-term effects of this shock on different cohorts.

other, unless one of them is assumed to have a nonlinear impact on change, their effect, let alone interactions, cannot be independently estimated from the data. To see why this is so, consider the following equation, where A, P, and C stand for age, period, and cohort, respectively: $A = P - C$; equivalently, $C = P - A$ or, alternatively, $P = A + C$. In other words, any of the three variables is a linear function of the others: Age can be directly deduced by knowing the year of the survey and the cohort in which the person belongs. Examining, for instance, respondents belonging in the 35-year-old cohort in 1987 is equivalent to examining those born in 1952. Evidently, we cannot isolate the impact of belonging in this cohort from the effect attributed to the role of age.

The usual way out is to constrain the impact of one of the three factors either to zero or to a nonlinear function. If, for instance, age is thought to be related to political participation in a curvilinear fashion, or cohort effects can be measured in a categorical way, a simple ordinary least squares (OLS) regression can capture adequately the independent effects of each confounding factor. To be sure, this leverage comes only at the cost of a more parsimonious specification, which assumes that period or cohort effects operate in the same way across cohorts with different years of birth. More recently, the use of Bayesian methods, generalized additive models, nonlinear regression models, discriminant analysis, or dynamic modeling of the effect of period shocks on different cohorts has helped examine the three factors independently, making only minimal assumptions about the way they affect the phenomenon of interest.

The second problem is that even when a typical regression model manages to produce point estimates for the effects of each of the three mechanisms, the standard errors will typically be unreliable. The reason for this is that the unit of analysis is usually an aggregate measure (percentage, mean values, etc.) of a group of individual responses taken from a sample of a targeted population. Thus, even in a case of a full-probability random sample, each cell represents an estimate, and thus, the uncertainty associated with this estimate needs to be also taken into account. The third problem is that tracing a cohort over a long period of time is also subject to nonrandom sampling error. As a given cohort becomes older, the number of observations representing it falls drastically. Importantly, the probability of survival may well be nonignorable to the trait of interest.

This last, less cited but probably more serious, problem of the ACP framework relates to its exploratory rather than explanatory character. Any ACP analysis is primarily useful as an effort to partition change into each of its components. However, since the year of birth and/or the year of survey cannot be logically treated as autonomous influences on any observed phenomenon of interest, unless some further examination of the causal processes that lead to this observation is made, any ACP account is only useful as suggestive evidence for what might be considered as the potential explanations for the underlying phenomenon. A more elaborate analysis, however, would need to model these processes and unpack the structural factors lurking behind each of the three terms. In statistical parlance, no analysis of this kind can account for the underlying process generating the data.

This last remark should not be interpreted as an argument against the use of cohort analysis altogether. That would mean throwing out the baby with the bathwater. In effect, cohort analysis, which started in the 1970s as a powerful tool that social scientists adopted from demographers to study patterns of social change, remains a creative strategy to disentangle interrelated sources of aggregate change. Although the difficulty of the ACP framework to separate these mechanisms creates limitations in the statistical analysis, it is also useful in showing researchers that social reality is often more complex than assumed by models, which fail to take into account all these competing mechanisms. Cohort analysis can be very useful either as a first explorative step accompanied by individual-level panel analysis or as a way to examine well-developed theories that mainly focus on one of the three mechanisms of change. Any such analysis needs to provide rigid theoretical argumentation and empirical evidence that would help unpack any of the three terms into substantive factors that operate through this process. In sum, cohort analysis can provide suggestive evidence about the dominant source of aggregate change when such exists or, as a minimum, make it clear that in various instances disentangling these effects is asking too much from the existing information that we have at hand.

Elias Dinas
European University Institute
San Domenico di Fiesole, Italy

See also Cross-Tabular Analysis; Electoral Behavior; Political Culture; Political Socialization

Further Readings

Fienberg, S. E., & Mason, W. M. (1984). *Cohort analysis in social science*. New York: Springer-Verlag.

Franklin, M. (2004). *Voter turnout and the dynamics of electoral competition in established democracies since 1945*. New York: Cambridge University Press.

Mannheim, K. (1952). The problem of generations. In K. Mannheim (Ed.), *Essay on the sociology of knowledge* (pp. 276–322). London: Routledge & Kegan Paul.

Mortimer, J. T., & Shanahan, M. J. (2003). *Handbook of the life course*. New York: Kluwer.

Tilley, J. R. (2003). Party identification in Britain: Does length of time in the electorate affect strength of partisanship? *British Journal of Political Science, 33*(2), 332–344.

Yang, Y. (2006). Bayesian inference for hierarchical age-period-cohort models of repeated cross-section survey data. *Sociological Methodology, 36,* 39–74.

Yang, Y. (2008). Social inequalities in happiness in the United States, 1972 to 2004: An age-period-cohort analysis. *American Sociological Review, 73,* 204–226.

Index

Entry titles and their page numbers are in **bold.**

International Encyclopedia of

POLITICAL SCIENCE

VOLUME 2

Editorial Board

International Encyclopedia of POLITICAL SCIENCE

VOLUME 2

EDITORS

BERTRAND BADIE
Sciences Po, Paris, France

DIRK BERG-SCHLOSSER
Philipps-Universität Marburg, Germany

LEONARDO MORLINO
LUISS Guido Carli, Italy

SAGE reference

Los Angeles | London | New Delhi
Singapore | Washington DC

IPSA AISP

Developed in partnership with the
International Political Science Association

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
33 Pekin Street #02-01
Far East Square
Singapore 048763

Publisher: Rolf A. Janke
Assistant to the Publisher: Michele Thompson
Developmental Editor: Diana E. Axelsen
Production Editor: Tracy Buyan
Reference Systems Manager: Leticia Gutierrez
Reference Systems Coordinator: Laura Notton
Copy Editor: QuADS Prepress (P) Ltd.
Typesetter: C&M Digitals (P) Ltd.
Proofreaders: Kate Peterson, Olivia Weber-Stenis
Indexer: Mary Mortensen
Cover Designer: Janet Kiesel
Marketing Manager: Kristi Ward

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

International encyclopedia of political science / edited by Bertrand Badie, Dirk Berg-Schlosser, Leonardo Morlino.

8 v. cm.
A SAGE Reference publication.
Includes bibliographical references and index.

ISBN 978-1-4129-5963-6 (cloth)

1. Political science—Encyclopedias. I. Badie, Bertrand. II. Berg-Schlosser, Dirk. III. Morlino, Leonardo

JA61.I58 2011
320.03—dc22 2011009801

11 12 13 14 15 10 9 8 7 6 5 4 3 2 1

Contents

2

List of Entries

Reader's Guide

Comparative Politics, Theory, and Methods

Anarchism
Anarchy
Breakdown of Political Systems
Cabinets
Censorship
Central Banks
Change, Institutional
Charisma
Citizenship
Collaboration
Comparative Methods
Comparative Politics
Competition, Political
Conditionality
Constitutional Engineering
Corporativism
Decentralization
Democracy, Types of
Discursive Institutionalism
Elites
Environmental Issues
Executive
Government
Historical Sociology
Human Rights, Comparative Perspectives
Hybrid Regimes
Institutionalization
Institutions and Institutionalism
Interest Groups
Irredentism
Labor Movement
Leadership
Legitimacy
Military Rule
Monarchy
Neo-Patrimonialism
Neo-Weberian State
Oligarchy
Path Dependence
Personalization of Politics
Pillarization
Political Integration
Political Science, International Institutionalization
Political Systems, Types
Politics of Language
Presidentialism
Prospect Theory
Qualitative Comparative Analysis
Referenda
Reform
Regime (Comparative Politics)
Regionalism
Regionalization
Representation
Republic
Republicanism
Responsibility
Responsiveness
Revolution
Rule of Law
Secession
Semipresidentialism
Separation of Powers
Social Movements
Socialist Systems
Stability
State
State, Virtual
Terrorist Groups
Totalitarian Regimes
Welfare Policies
Welfare State

Case and Area Studies

Area Studies
Authoritarian Regimes

Subnational Governments

Democracy and Democratization

Decision Making in Democracies

Epistemological Foundations

Gender and Race/Ethnicity

International Relations

Political Science, International
Institutionalization
Postmodernism in International Relations
Psychological Explanations of International Politics
Realism in International Relations
Superpower

Peace, War, and Conflict Resolution

Alliances
Arms Race
Bilateralism
Bipolarity and Multipolarity
Civil War
Collective Security
Conflict Resolution
Conflicts
Détente
Diplomacy
Disarmament
Domestic Politics and International Relations
Empire
Foreign Policy Analysis
Genocide
Imperialism
Intervention
Intervention, Humanitarian
Judicial Decision Making
Judicialization of International Relations
Mediation in International Relations
Militias
Multilateralism
National Interest
Natural Resources
Neutrality
Pacifism
Participation, Contentious
Peace
Peacekeeping Positive Peace
Power and International Politics
Preemptive War
Psychological Explanations of International Politics
Sanctions
Secession
Security and Defense Policy
Security Cooperation
Security Dilemma
Sovereignty
Strategic (Security) Studies
Superpower
Territory
Terrorism, International
Transatlantic Relations
Unilateralism
United Nations
Violence
War and Peace
Warlords
Westphalian Ideal State
World Systems Theory

Political Economy

Capitalism
Central Banks
Class, Social
Cost–Benefit Analysis
Economic Policy
Economic Statecraft
Economic Theories of Politics
Foreign Aid and Development
Inequality, Economic
International Monetary Fund (IMF)
International Political Economy
Labor Movement
Market Economy
Market Failure
Monetary Relations
Multilateralism
Multinational Corporations (MNCs)
Nongovernmental Organizations (NGOs)
Policy, Employment
Political Economy
Privatization
Property
Protectionism
Public Budgeting
Public Employment
Public Goods
Redistribution
Social Stratification
Sustainable Development
Tax Policy
Trade Liberalization
Traditional Rule
Tragedy of the Commons
Transaction Costs
Transformation, Economic

Political Parties

Political Philosophy/Theory

Equality and Inequality

Formal and Positive Theory

Theorists

Political Sociology

Public Policy and Public Administration

Public Policy

Policy Formulation
Policy Framing
Policy Instruments
Policy Learning
Policy Network
Policy Process, Models of
Prospect Theory
Reorganization
Risk and Public Policy
Self-Regulation
Soft Law
Stages Model of Policy Making
Think Tanks
Tragedy of the Commons
Transaction Costs

Public Administration

Administration
Administration Theory
Audit Society
Auditing
Autonomy, Administrative
Budgeting, Rational Models
Bureaucracy
Bureaucracy, Rational Choice Models
Bureaucracy, Street-Level
Civil Service
Corruption, Administrative
Effectiveness, Bureaucratic
Governance
Governance, Administration Policies
Governance, Informal
Governance, Multilevel
Governance, Urban
Governance Networks
Groupthink
Health Policy
Intelligence
Pay for Performance
Performance
Performance Management
Planning
Police
Politicization of Bureaucracy
Politicization of Civil Service
Public Budgeting
Public Employment
Public Goods
Public Office, Rewards
Regulation
Representative Bureaucracy
Responsiveness of Bureaucracy
Secret Services
Security Apparatus

Qualitative Methods

Analytic Narratives: Applications
Analytic Narratives: The Method
Configurational Comparative Methods
Data, Textual
Discourse Analysis
Ethnographic Methods
Evaluation Research
Fuzzy-Set Analysis
Grounded Theory
Hermeneutics
Interviewing
Interviews, Elite
Interviews, Expert
Mixed Methods
Network Analysis
Participant Observation
Process Tracing
Qualitative Comparative Analysis
Quantitative Versus Qualitative Methods
Thick Description
Triangulation

Quantitative Methods

Aggregate Data Analysis
Analysis of Variance
Boolean Algebra
Categorical Response Data
Censored and Truncated Data
Cohort Analysis
Correlation
Correspondence Analysis
Cross-National Surveys
Cross-Tabular Analysis
Data, Archival
Data, Missing
Data, Spatial
Data Analysis, Exploratory
Data Visualization
Event Counts
Event History Analysis
Experiments, Field

Religion

COLD WAR

See Balance of Power; Bipolarity and Multipolarity; Superpower; Totalitarian Regimes; Transatlantic Relations; War

COLLABORATION

Collaboration is a process in which the state engages with business and civil society to shape policy agendas, and to design and implement policy programs. Sometimes this is interpreted as meaning that the public policy process reflects the principles of deliberative democracy, in which actors engage in a rational debate with the intention of seeking a consensus solution. Other literatures place more emphasis on collaboration arising from resource interdependencies and complex problems that require a multisectoral response. A more critical understanding of collaboration is as one of the mechanisms through which the state subjugates populations via new mechanisms of self-governance.

Despite definitional differences, there is general agreement that collaboration is different from cooperation and coordination. All, however, are points on a continuum from organizational independence to organizational integration. Cooperation involves sharing information or expertise in a context where organizations remain autonomous from each other, for example, in client referrals in social welfare networks. Coordination involves a process of independent organizations introducing a degree of integration by undertaking mutual adjustment to produce a better overall outcome. This takes place when organizations agree to redesign their procedures so that, for example, clients can use a single point of contact to access services from a variety of agencies. Collaboration is the next step in the continuum. It involves organizations recognizing that they are interdependent and creating a greater degree of integration of procedures, policies, and structures. Such collaboration might involve the creation of an overarching board to make strategic decisions that are binding on the constituent organizations.

Collaboration is often regarded as an activity that takes place at the subnational level, and there is a large literature on partnership and other forms of collaboration at city or neighborhood level, for example, in urban regeneration and sustainability. However, it would be incorrect to privilege these spatial scales. In an era of multilevel and global governance, it is important that the discussion of collaboration also take into account interactions at national and transnational level, as well as interactions between these and other levels of governance. This is because collaborative mechanisms engaging government with civil society and business have become the preferred delivery mechanisms for many aspects of European and national policy, including those related to economic restructuring, social inclusion, environmental sustainability, and the knowledge society. These policy initiatives stimulate activity at the subnational level and in some cases (as in a number of European Commission initiatives) become the instrument through which specific policy intentions are expected to be realized. In the field of international relations, the creation by states of instruments of global governance can be regarded as the outcome of a process that is intended to lead to greater collaboration, for example, through the regulation of standards or economic, environmental, or social policies. Collaboration, therefore, involves both horizontal and vertical forms of interorganizational engagement.

How Does Collaboration Relate to Other Concepts?

The concept of collaboration is closely related to other concepts in common usage. The first term to consider is *governance*. In political science and public policy, this typically refers to a changing form of state rule through which government attempts to engage other actors in its political projects through cultural rather than coercive mechanisms. However, the extent to which actors from business and civil society are able or willing to engage in collaborative activity with government will be influenced by the nature of the state and its historical relationships. So the experience of the United States with its tradition of "small government" and extensive private provision of public services will be very different from that of Western Europe, where the state has played a much more prominent role in service provision. The emergence of multilevel governance is especially a recent consequence of European integration, although decentralization and different levels of government are a part of traditions inside and outside Europe.

Second, collaboration is broadly synonymous with the concept of "partnership." This concept is used in some English-speaking countries (e.g., the United Kingdom [UK], Australia, New Zealand) to refer to the quasi-formal organizational arrangements that provide an envelope for interaction between state, civil society, and business on particular policy questions. In these countries, "public–private partnership" typically refers to contractual agreements between government and business related to the provision of infrastructure or services, whereas in Europe and the United States, it tends to encapsulate both this specific meaning and the wider Anglo idea of partnership. For U.S. scholars, "collaborative public management" has a meaning synonymous with the idea of "partnership" in the UK.

The third concept with which collaboration connects is that of "networks," which develops the older idea of policy networks. Here, the emphasis is on an institutionalized pattern of relationships between autonomous actors in a particular policy realm. Negotiated agreements between actors could be understood as collaboration to the extent that the actors agree to forgo some of their autonomy in order to realize collective benefits.

Why Is Collaboration Important Today?

The contemporary importance of this idea arises from the changing role of the state in the late 20th and early 21st century. The argument is that the state lacks the necessary capacity and resources to deliver public policy goals in advanced capitalist economies and chooses to fill this gap by involving others sectors more into functions that would—under traditional public administration—have been reserved to the state and its functionaries. Collaboration is therefore part of the new public management movement, which includes the view that public services could be delivered more efficiently and effectively by external providers—including the widespread fascination with the involvement of the private sector.

A related explanation is that neoliberal ideology, especially in social welfare and former socialist states, has resulted in a fragmentation of government due to privatization, contracting out, the creation and hiving off of separate agencies from large public bureaucracies, and other new public management initiatives. Collaboration arises from the need to retain a capacity to integrate across this distributed organizational landscape so that policy intentions can be realized and steering is possible. Such integration occurs at a number of levels. At the discursive level, there are powerful ideas—for example, about "partnership"—which mobilize actors around collaborative projects despite the differences between them. At the institutional level, collaboration involves a process of rule creation where there is greater flexibility for actors than in representative government. And at the behavioral level, boundary spanners and network managers are valued because of their capacity to build coalitions between organizations and groupings in an ambiguous and changing environment.

What Are the Democratic Implications of Collaboration?

From a progressive perspective, collaboration offers a way of developing participative forms of governance that bypass, or at least complement, representative democracy. It facilitates the engagement of a wider range of actors in the policy process and potentially enables them to contribute throughout the process rather than just being consulted at the start and excluded from the black box of decision making. So collaboration can be treated as synonymous with deliberative and participative democracy, and the opening up of the state to other interests (and especially those of citizens and civil society actors). This progressive perspective is reflected in the normative tone that can be found in the literature.

A more critical reading positions collaboration in the light of Foucauldian governmentality, where close engagement with the state introduces a regime of discipline on participants. In this process, they become part of the mechanisms of governance in a complex society, taking on themselves the concerns of the state and self-regulating their actions in this way. Thus, citizens involved with the state in regenerating disadvantaged communities take on the mantle of the state's discourse of disadvantage and the performance management systems that go along with delivering public policy. This constrains their capacity to think and act outside this discourse in ways that might serve their interests better but go against the state. On this basis, citizens might adopt a political strategy of exit rather than remain trapped by the false reality of collaboration.

This discussion reminds us that collaboration—for all its progressive overtones—is about political action. The prevailing discourse of collaboration conceals the darker side in which decisions are removed from the arena of representative government with its public interest safeguards and located instead in less transparent institutional contexts whose rules in use are evolving and more malleable. If we agree that the relations of power between state, business, and civil society are structured to give advantage to the former two actors, and thus biased against civil society, then collaboration is likely to do little to redress these more fundamental imbalances. The challenge for academics analyzing these new arrangements is to investigate beneath the epiphenomenon of political discourse and understand the nature and impact of collaboration on the relations between actors in historically embedded contexts.

Chris Skelcher
University of Birmingham
Birmingham, United Kingdom

See also Deliberative Policy Making; Governance; Networks; New Public Management; Policy Analysis; Policy Instruments; Risk and Public Policy

Further Readings

Agranoff, R., & McGuire, M. (2003). *Collaborative public management*. Washington, DC: Georgetown University Press.

Koppenjan, J. F. M., & Klijn, E.-H. (2004). *Managing uncertainties in networks: A network approach to problem solving and decision-making*. London: Routledge.

Sullivan, H., & Skelcher, C. (2002). *Working across boundaries: Collaboration in public services*. Basingstoke, UK: Palgrave Macmillan.

Collective Security

A collective security system is one in which each state in the system accepts that the security of one is the concern of all and agrees to join in a collective response to threats to, and breaches of, the peace. A weaker version posits a system that commits governments to develop and enforce broadly accepted international rules in the area of international peace and security and to do so through collective action legitimized by international institutions.

The idea of collective security goes back at least to the European peace plans of the 18th century and gained ground in the post–World War I period as international society sought to restrict the previously wide-ranging right of states to resort to war as an instrument of state policy, first in the League of Nations and then in the United Nations (UN). The UN did not constitute a pure collective security system (e.g., in relation to the veto given to permanent members of the Security Council); but it contained powerful elements of collective security, above all in terms of the clear prohibition of aggressive force in Article 2(4) of the UN Charter, and in the far-reaching responsibility of the Security Council for the maintenance of international peace and security, including the authorization of mandatory sanctions and military action. The end of the Cold War appeared to many to open new opportunities for collective security, with an increasing number of enforcement resolutions adopted under Chapter VII and a very significant expansion in the number and coercive character of UN peacekeeping operation and of UN-established international administrations.

The Nature of Collective Security

Collective security raises five sets of questions about its nature, with respect to the participants in collective security arrangements, the kind of security provided, and the forms of collective action to be taken.

Security for Whom?

The traditional conception of collective security was intended to strengthen the rights of states to independence and to reinforce an international legal order built around the concepts of sovereignty and nonintervention. On the Wilsonian view, one of its chief attractions was that, in contrast to the balance of power, it guaranteed the independence of all states, including small and weak states. Yet the stress on reinforcing the rights of states and on maintaining the sanctity of established borders against forcible change has given rise to one of the most enduring dilemmas: how to accommodate change and how to prevent a collective security organization from becoming an instrument for maintaining the status quo. The liberal assumption of a shared interest in the peaceful maintenance of the status quo has been repeatedly challenged.

What Kind of Security?

Second, what kind of security is embodied in the phrase *collective security*? As it developed in the late 19th and early 20th centuries, collective security was conceived as a response to the dangers of formal interstate violence and, in particular, to the problem of the aggressive use of force by states. Faced with the united opposition of the international community, states would come to accept that aggression simply could not pay. Although not as novel as sometimes claimed, the post–Cold War period has seen an enormous literature on the changing character of security and the changing dynamics of the global security landscape: the easing of major power rivalry and the emergence of a wide range of new security challenges connected with civil wars, domestic social conflicts, ethnic strife, refugee crises, humanitarian disasters, and transnational terrorist threats. In addition, many have argued that much greater priority should be given to human security, rather than the security of states or of regimes. Such moves illustrate the politically contested character of the concept of security, emphasized, in particular, by critical security theory. Collective security is sometimes understood analytically as a problem of capturing a well-understood and broadly shared interest (perhaps by viewing security as a global public good) in the face of the twin problems of defection and free riding. Clearly, these problems are formidable in a world in which states and government leaders are faced with powerful incentives to protect their immediate short-term interests. But such rationalist logic underplays the challenge of the essentially contested nature of security. Very different historical circumstances and divergent values mean that there is rarely an easy answer to the question of whose security is to be upheld or against which threats that security is to be promoted. Together with the deeper intervention required to deal with many new security challenges, this in turn increases the problems of legitimacy as well as the difficulties of securing the willingness of states to commit armed forces to conflicts that are often seen as marginal to core foreign policy interests.

Which Collectivity?

Third, which collectivity is involved in "collective security"? There have always been strong arguments for the broadest possible membership of a collective security system: to ensure that the power of the collectivity is sufficient to deter aggression and, if necessary, to enforce its decisions against all states, and to reduce the danger that collective security will merely provide a framework within which power political competition and alliance politics are played out under a different guise. On the other hand, there have also been repeated arguments that an effective collective security system requires leadership and that an effective collectivity will consist of a smaller group of like-minded states with the effective (as opposed to theoretical) power to enforce their decisions. Hence, there have been recurrent arguments that regionally based collective security systems are most likely to prove effective, and the post–Cold War period has seen an expansion of the role of regional security organizations.

What Form of Collectivity?

Fourth, what form of collectivity? It is important to disentangle two approaches to thinking about collective security systems. In one view, the dominant one in the recent past, collective security has been seen as a means of enforcing order between independent political communities, of achieving a degree of centralization that does not radically threaten the independence and autonomy of states. An alternative, and historically deep-rooted, conception has viewed moves toward the collective management of armed force as part of a broader process of reorganizing the political system and moving beyond the state system toward more centralized or federal forms of global political order.

The other issue here concerns the dilemma of preponderant power. In theory, collective security offers the purest solution to the dilemma of preponderant power. Inequality is not to be feared, opposed, or balanced against but is, instead, to be harnessed to the legitimate collective purposes of the international community. In practice, the situation is more complex. On the one hand, the veto within the UN Security Council reflects the reality of a power distribution in which attempts to coerce any of the major powers of the system could be achieved only at great risk and high cost. On the other hand, unable to command substantial military forces of its own in the ways envisaged in the Charter, UN enforcement action has operated by means of authorizing the use of force by member states, especially those with the capacity to deploy effective military power. Such a situation is always likely to create problems of effective delegation and control and to increase the risk that a system aimed at collective security will in fact become one of selective security.

What Forms of Collective Action?

Finally, what forms of collective action are envisaged in a collective security system? Collective security involves a shared acceptance that a breach of the peace threatens the interests of all states. It also involves a shared willingness to act effectively to enforce the law and to protect the interests of the international community. Enforcement has very often been seen as critical, but not always. Much of 19th-century liberal thought, for example, believed fervently that the clear elaboration of international law would be crucial. Others believed in the power of enlightened international opinion. The arguments for effective enforcement grow through the century and are, of course, given particular force by the catastrophe of World War I. At the same time, however, these older liberal views never entirely faded, and there remained the hope that the paradox of fighting a war to ensure peace could be avoided.

It should also be noted that the emphasis on deterrence and enforcement places collective security analytically close to mainstream realist writing on alliances and the balance of power. This can be contrasted with constructivist writing on cooperative security and on security communities. It is the possible emergence of a situation in which cooperation goes beyond instrumental calculation and in which the use of force declines as a tool of statecraft that opens the door to constructivist theories. Such theories highlight the importance of historically constructed interests and identities, of learning and ideational forces, and of normative and institutional structures within which state interests are constructed and redefined. Instead of focusing solely on material incentives, constructivists claim that understanding intersubjective structures allows us to trace the ways in which interests and identities change over time and new forms of security cooperation and community can emerge.

Liberals argue that, although there is no collective security *system*, the collective *element* in security management has increased and even major states need multilateral security institutions both to share the material and political burdens of security management and to gain the authority and legitimacy that the possession of crude power alone can never secure. Realists continue to stress that peace is not indivisible, that states and their citizens are often unwilling to bear the costs of collective action in conflicts in which their direct interests are only weakly engaged, and that large parts of the global security system continue to be shaped by the unilateral pursuit of state interest and by practices of both the balance of power and hierarchical power.

Andrew Hurrell
Oxford University
Oxford, United Kingdom

See also Balance of Power; Civil War; Intervention, Humanitarian; Liberalism in International Relations; Peacekeeping; Security Cooperation; State Failure; War and Peace

Further Readings

Adler, E., & Barnett, M. (Eds.). (1998). *Security communities.* Cambridge, UK: Cambridge University Press.

Bellamy, A. J., & William, P. D. (2005). Who's keeping the peace? Regionalization and contemporary security operations. *International Security, 29*(4), 157–195.

Claude, I. L. (1962). *Power and international relations.* New York: Random House.

Downs, G. (Ed.). (1994). *Collective security beyond the Cold War.* Ann Arbor: University of Michigan Press.

Hurrell, A. (2007). *On global order: Power, values and the constitution of international society*. Oxford, UK: Oxford University Press.

Krahmann, E. (2008). Security: Collective good or commodity? *European Journal of International Relations, 14*(3), 379–404.

Lowe, V., Roberts, A., Welsh, J., & Zaum, D. (Eds.). (2008). *The United Nations Security Council and war: The evolution of thought and practice since 1945*. Oxford, UK: Oxford University Press.

COLONIALISM

Colonialism is usually understood as a political doctrine promoting and justifying the exploitation by a colonizing power of a territory under its control either for its own benefit or for the benefit of the colons settled in this territory. In this sense, colonialism refers mainly to the unequal relationships developed between European colonizers and their respective colonial empires. This conventional definition has been considered too restrictive by some scholars who, in the past 20 years, have stretched the notion. Colonialism has come to include many different kinds of unequal power relationships between two countries (e.g., between Israel and Palestine) and between the West and the world (the concept of colonialism replacing to some extent that of imperialism). Colonialism can also refer to unequal relationships between a dominant majority group and a minority group that is an indigenous group or considered to be not autochthonous (internal colonialism). Colonialism has also been used in association with larger modern political and economic processes such as the economic world system since the 16th century or, more generally, with a vision of European "modernity." The polysemy of colonialism is largely due to the renewed interest, since the 1980s, in the colonial and postcolonial periods among literary critics, historians, anthropologists, and political scientists. In this entry, colonialism is analyzed only as a set of complex, unequal, and past relationships linking European colonizing powers to colonies (for other meanings, see the entry on postcolonialism). It is, thus, useful to dissociate colonialism as a political doctrine forged by those writing during the colonial period from colonialism as a paradigm reexamined by scholars since the 1980s. In the last section of this entry, the way postcolonial thinkers have portrayed colonialism is briefly explored before other approaches examining colonialism from a more political, social, and economic angle are examined.

Colonialism as a Political Doctrine

The word *colonialism* is created from the words *colonial* and *colony*, which have a longer history. English and French terms for *colonial* and *colony* derive from *colonus*, the Latin word for "farmer." The founding of colonies was one of the strategies the ancient Romans employed in establishing their empire. During the Renaissance and the 16th-century European expansion, the words *colon* and *colony* gradually took on their current meaning: The word *colon* referred to a person living in a colony, as opposed to an inhabitant of a European colonial power; the world *colony* designated a territory dominated and administrated by a foreign power or a group of settlers; the verb *to colonize* gradually came to mean "to conquer a territory." At different periods but mainly in the second half of the 18th century, the words *colonial*, *colonize*, and *colonization* appeared in English and were then translated into French.

The term *colonialism* obviously derived from *colonial*, but it appeared later, in the framework of 19th-century imperialism. It first appeared in English around the middle of the 19th century and was used to mean practices or idioms peculiar to, or characteristic of, a colony. In 1886, it was used to mean the colonial system or principle, thus referring to colonialism as a systematic and wide-ranging phenomenon. In France, it also followed the pace of overseas expansion. The word *colonist*, common after the conquest of Algeria in 1830, referred specifically to a partisan of the colonization of Algeria, while *anticolonist* referred to opponents of this process of colonization. This specific use did not last, however, and *colonialism/anticolonialism* came to replace the terms *colonist/anticolonist* in the early 20th century.

It is thus not really surprising that colonialism became a political doctrine promoted by theorists and defended by interest groups only in the 19th century, during the heyday of imperialism. Colonialism can hardly be understood without the element of self-legitimation inherent in it. The most powerful form of self-legitimating was the colonizer's claim of improving the conquered country and bringing the fruits of progress and modernity to the subject peoples. All European powers claimed to pursue a civilizing project in their colonies, but they used different terminologies to this end. In the 19th century, it was called *improvement*, *betterment*, or *moral and material progress*. All these terms were then subsumed under the term *civilizing mission*, which became the imperial ideology and official doctrine from the late 19th century onward. Parallel to this historical trend, there was an increasing feeling of racial superiority in European nations, along with the development of a positivist approach in the natural and human sciences associated with an increasing obsession with classifying people, plants, and animals. Some Western philosophers, academics, writers, and politicians developed a new vision of the world. There were continents with history and those without history (Friedrich Hegel, 1770–1831); there were superior and inferior races to colonize (Jules Ferry, 1832–1893), fortunate and less fortunate races to educate (Rudyard Kipling, 1865–1936), inferior and superior languages (Ernest Renan, 1823–1892), and a primitive or prelogical mode of thought and a logical one (Lucien Levy-Bruhl, 1857–1939).

In the 19th century, however, colonialism as a political ideology was not yet a unifying body of knowledge. In the British Empire, the aim for moral improvement and material betterment of society became common among a generation of administrators guided by the 18th-century legal reforms of Warren Hastings (1732–1818), Lord Cornwallis (1738–1805), and Jeremy Bentham (1748–1832). Gradually, British historiography started to describe colonization as a history of progress. This is well exemplified by James Mill's *History of British India*, an early and influential piece of colonial historiography that was published in 1818. In contrast, a specific doctrine of modern colonization emerged in France only in the early years of the Third Republic (1871–1940). Following Germany's defeat of France in 1870–1871 and France's loss of the two provinces of Alsace and Lorraine, some disparate groups developed a common set of arguments to justify overseas expansion. In 1874, the economist Paul Leroy-Beaulieu (1843–1916) advanced the view in his famous book *De la colonisation chez les peuples modernes* (Colonization among modern nations) that investing in colonies was the best business for an old and rich country such as France. The most influential Republican leaders, Léon Gambetta (1838–1882) and Jules Ferry, also defended colonial expansion as a way to recover from the events of the 1870s. While colonization initially divided monarchists, republicans, and radicals in France, no political parties in Europe opposed colonization prior to World War I.

The making of 20th-century colonialism was a process driven by a set of intellectuals, writers, and academics and by a new administrative machinery (the colonial office) assisted by experts (ethnographers, ethnologists, and later on anthropologists and sociologists) and increasingly after World War II by technicians (medical officers, educationists, town planners, welfarists, etc.). While the idea of progress was very common in justifications of colonial conquest in the 19th century, a considerable debate emerged in the early colonial days of administration concerning what kind of colonial rule was desirable. Assimilation, association, and indirect rule are generally considered to be the three major colonial doctrines that emerged between the end of the 19th century and the 1920s. This debate became most intense in France and Germany; it was more restricted in the United Kingdom, where there was general agreement that indirect rule should be the underlying philosophy of the system of government. The doctrine of assimilation was based on the heritage of the French Revolution and the belief in the equality of all people irrespective of their racial origin or cultural background. Colonized people had to prove themselves worthy of assimilation by demonstrating to the authorities that they had the attributes needed for citizenship. Practically, however, access to full rights of French citizenship was highly restricted, mainly to the inhabitants of four communes in Senegal (Dakar, Gorée, Rufisque, and Saint-Louis) and to a limited number of educated people (*évolués*) before World War II. In contrast, the doctrine of indirect rule was inspired by the idea that Europeans and Africans were culturally distinct, and indigenous political institutions were necessary for the purposes of local government. This was conceptualized by Lord Frederick Lugard, Governor General of Nigeria from 1914 to 1919, in his *Dual Mandate in Tropical Africa*, published in 1922. Rather than preserving the precolonial polity, indirect rule largely modified the public authority of traditional chiefs to ensure the exploitation of the country through means such as the introduction of taxes and forced labor and the modification of customary law. The differences in colonial doctrines should not be overstated. For instance, association was mainly another word for indirect rule in French colonial theory. In the administrative offices within the colonies themselves, colonial theory was less important than in European capitals; most of the time, colonial administration used indirect rule without acknowledging it as the cheapest way to govern colonies.

The participation of soldiers from European colonies in the two World Wars, the multiplication of colonial exhibitions all over Europe, the increasing number of colonial magazines, and the teaching of the history and geography of the empires in schools generated popular support for European colonization between the two wars. The legitimacy of colonialism was little contested at the time. Anticolonialism was limited to a few writers and to the extreme left (although the Communist Party in general was not really anticolonial when it was in government). It was only during and after World War II, with the rise of nationalism in Asian and African countries, that colonialism became radically contested both in colonizing and in colonized countries. Providing different rights to different peoples on the basis of their racial classification became illegitimate after the defeat of the Nazis and the rise of a bipolar world. In addition to the revolutionary anticolonialism of the far left, there was a broader moral opposition to violence arising in the colonial context, such as that in Madagascar, Kenya, and Algeria.

The various European colonial powers reacted differently to the new postwar order. Portugal decided not to concede independence or freedom to its African colonies, while other European governments granted new political and social rights and tried to reshape a more legitimate colonialism under the new ideology of development through technology, public health, and economic and local government reforms. Subjects were transformed into citizens, and they began to demand the same social and political rights as those in the European colonizing countries. The cost of this new dispensation, however, was too high for postwar European budgets; these economic concerns, together with the rise of nationalism, convinced European leaders to give up formal political ties with their former colonies. In less than 15 years, colonialism ceased to be a legitimate political doctrine and ultimately disappeared in the early 1960s. It survived only in Portugal, in Portuguese colonies and in settler colonies, where it was described sometimes as internal colonialism (as, e.g., in South Africa or Rhodesia). Elsewhere, the notion of colonialism and anticolonialism was replaced by neocolonialism, a neologism used to describe, and more often to denounce, the conditions under which former colonies continued to serve the economic, political, military, and other interests of powerful, mostly Western, countries. The concept of neocolonialism was popularized by the new independent leader of Ghana, Kwame Nkrumah, in his 1964 book *Neo-Colonialism, the Last Stage of Imperialism.*

Colonialism as a Contested Paradigm

When colonialism was a highly contested political issue in the 1950s, the French sociologist Georges Balandier wrote a pioneering article on the colonial situation (1951), in which he explained the necessity of understanding colonization as both a social phenomenon and a specific historical process. Balandier's recommendation was not followed, however, as researchers in the 1960s and 1970s were no longer interested in a political ideology that had been banished from the realm of legitimate forms of political organization. As mentioned by Frederick Cooper, there is something strange about the writing of colonialism: Scholarly interest in colonialism arose when colonial empires had already lost their international legitimacy. This rediscovery is partly explained by the rise, in academic institutions of the English-speaking world, of what has been labeled as postcolonial studies.

During the first decades after independence, the ideology of progress and modernization (then perceived to be Western) was largely shared by the political and intellectual elites of former colonies. However, the various projects of modernization of the state and the economy implemented after independence were seen by critics as inefficient and socially unjust. It is against this background that a critique of colonialism reemerged within a group of intellectuals from the former colonial world, who were mostly educated in Western academic institutions. Edward Saïd's book *Orientalism*, published in 1978, is often considered to be a starting point for the analysis of colonialism as a specific configuration of knowledge and power. Influenced by the work of Michel Foucault, Saïd argues that the intellectual construction of the Orient by Western scholars has served as an implicit justification for the colonial and imperial ambitions of European nations and the United States. Valentin Mudimbe, 10 years later, proposed a similar textual and cultural reading of the colonial domination of Africa, asserting that Africa was invented by a colonial and anthropological discourse and this process was a way for Europeans to build their own identity. Also important in the critique of the European colonial ideology was the initiative launched by a group of Indian historians such as Ranajit Guha, Partha Chatterjee, Gyanendra Pandey, and Dipesh Chakrabarty, who contested the nationalist historiography of India. In 1982, they initiated a series of edited books titled *Subaltern Studies*, which focused on the subaltern groups (or oppressed people) of colonial India. The central notion of *Subaltern Studies* is the notion of "agency," which includes autonomy of the subaltern in the political arena and a consciousness of self that is not controlled by the Western elites and their nationalist counterparts, who have supposedly adopted the values of their colonizers. This agency can be particularly understood in the words and knowledge of subaltern people.

The major aim of postcolonial studies was to deconstruct a Western epistemology associated with colonialism that reified the non-Western world in an ahistorical moment. The result was a critique of the civilizing-mission discourses; the works of Aimé Césaire, *Discourse on Colonialism* (1953); Albert Memmi, *The Colonizer and the Colonized* (1957); and Franz Fanon, *The Wretched of the Earth* (1961) received new attention from the 1980s onward as texts that pioneered the denunciation of the pernicious cultural and psychological effects of colonialism. This rediscovery has largely popularized the idea that colonized people have no other story than that of colonial oppression.

According to the historian Frederick Cooper, a significant part of postcolonial thought has taken colonial studies out of its historical context, treating colonialism abstractly, generically, as located somewhere between 1492 and the 1970s, and as something to be juxtaposed with an equally flat vision of European "modernity." Moreover, colonial societies should not be understood merely in terms of a "European versus indigenous" dichotomy, and the actions of former colonized people should not be reduced to merely resistance to colonialism or collaboration with the colonial order. Similarly, when colonialism is portrayed as a power-demarcating and -racializing space, the complexities of the colonial situation are ignored or one tends to neglect the internal divisions within European and indigenous communities alike.

The analysis of the relationship between colonialism and biomedicine can be used as an example to explain the capacity of colonialism to produce both a contradictory and a generic discourse on colonized people. According to Megan Vaughan, who traced the emergence of various medical discourses in Africa, there was, on the one hand, a long tradition, elaborated by Christian missionaries, that viewed the "primitiveness" of African societies as a factor predisposing them to certain diseases. On the other hand, research on the relationship between colonialism and biomedicine from the interwar period indicated that Africans got sick because they forgot who they were: Urbanization, industrialization, and deculturation were seen as factors leading to insanity, sexually transmitted diseases, or leprosy. Taken together, these contradictory viewpoints show, however, a fundamental difference in the emergence of biomedical discourse in Europe: Africans were always conceived of as members of a collectivity—as colonial people—and beyond that as members of collectivities—in the form of tribes or cultural groups—and this led to different views in tracing the relationship between colonialism and biomedicine.

This reification of groups by colonialism has been contested by social historians and anthropologists. Groups and categories under colonial rule (chiefs, the educated elite, traders, peasants, workers, and also women, youth, and elders) pursued their own agendas, defended their own interests, and consequently changed the limits of subordination within the colonial system. In this sense, opposing too markedly the "bourgeoisie" and the "working class" or the "elite" and the "people" includes the risk of reproducing the old colonial dichotomy as well as creating new ones (opposing the modern to the traditional). A large part of African and Indian historiography in the past 3 decades has tried to overcome this vision present in colonialism by looking at the agency of individuals and groups, the contingent nature of their connections and networks, and their respective roles in shaping power relations. Actually, colonialism did not produce only differences between colonizers and the colonized. It also produced individuals with varying investments in a range of identities, sexual identities, class identities, religious identities, and ethnic identities. Similarly, colonialism not only led to conflicts between colonizers and the colonized, it was also decisive in shaping new roles, divisions, and conflicts between elders and youth, men and women, and migrants and urban dwellers.

The understanding of conflicting colonial repertoires helps explain why colonialism was not pervasively efficacious in implementing modernization. According to Nicholas Thomas (1994), "colonialism was not a unitary project but a fractured one, riddled with contradictions and exhausted as much by its internal debates as by the resistance of the colonized" (p. 51). Colonized people were able to turn the discourse of modernization into a language of claims concerning the obligations of colonial powers. For instance, when development emerged as a colonial project with aid from the former colonial power, it did so in the face of serious objective challenges from the West Indies and Africa in the 1930s and the 1940s.

The various forms of present-day oppressions in some former colonies have sometimes been regarded as a legacy of colonialism. For instance, Mahmood Mamdani (1996) argues that colonial rule created a "bifurcated state" that distinguished between a despotic tribal power in rural areas and a democratic civil society based in towns and cities. According to Mamdani, this duality of power has been a major obstacle to democratization in postindependence Africa. However, historians have shown that colonial power was limited by tribal chiefs' obligation to ensure the well-being of their community to maintain the legitimacy on which colonial authorities depended. As mentioned by Thomas Spear (2003), colonial dependence on chieftaincy often limited colonial power as much as facilitating it. The reification by colonialism of categories of citizen/subject and urban/rural overstates the role of urban and elite power and undervalues the strategic place of rural constituencies and the importance of urban–rural links in the making of African politics. Eventually, drawing a direct causal connection between the indirect rule of the 1920s and 1930s and the politics of authoritarianism and ethnicity in the 1980s and 1990s fails to see what lies in between, especially the explosion of citizenship in the final years of colonial rule.

Historians, anthropologists, and political scientists interested in understanding the everyday practices of colonialism have changed their framework of analysis to take account of these issues. They today insist on the necessity of looking simultaneously at the colonizing countries and the colonies and the process by which each entity affected the political transformation of the whole empire. This marks an important break with former imperial historiography, which had long treated colonies as marginal to a history that remained national or as a projection of national culture and power. Empirical evidence now suggests that new forms of administration, town planning, architecture, policing, and medical practice, to mention just a few examples, were not only a one-way imposition of norms coming from the metropolis but were also coproduced in the colonies and contribute to larger debates about the nature of colonialism and its impact. Colonialism profoundly shaped both colonized and colonizers' society.

Laurent Fourchard
Institut d'Etudes Politiques de Bordeaux
Pessac, France

See also Imperialism; Postcolonialism

Further Readings

Balandier, G. (1951). La situation coloniale: Approche théorique [The colonial position: Theoretical approach]. *Cahiers internationaux de sociologie* [International Sociology Notebooks], *11,* 44–79.

Cooper, F. (2005). *Colonialism in question: Theory, knowledge, history*. Berkeley: University of California Press.

Cooper, F., & Stoler, A. L. (Eds.). (1997). *Tensions of empires: Colonial cultures in a bourgeois world*. Berkeley: University of California Press.
Fischer-Tiné, H., & Mann, M. (Eds.). (2004). *Colonialism as civilising mission*. London: Wimbledon.
Mamdani, M. (1996). *Citizen and subject: Contemporary Africa and the legacy of late colonialism*. Princeton, NJ: Princeton University Press.
Mudimbe, V. I. (1988). *The invention of Africa: Gnosis, philosophy and the order of knowledge*. Bloomington: Indiana University Press.
Pouchepadass, J. (2007). Le projet critique des postcolonial studies entre hier et demain [The critical postcolonial studies project between yesterday and tomorrow]. In M. C. Smouts (Eds.), *La situation postcoloniale: Les postcolonial studies dans le débat français* [The postcolonial position: The postcolonial studies in the French debate] (pp. 173–218). Paris: Presses de la Fondation Nationale des Sciences Politiques.
Said, E. (1978). *Orientalism*. New York: Pantheon.
Spear, T. (2003). Neo-traditionalism and the limits of invention in British colonial Africa. *Journal of African History, 44*, 29–50.
Thomas, N. (1994). *Colonialism's culture: Anthropology, travel and government*. Cambridge, UK: Polity Press.
Vaughan, M. (1991). *Curing their ills: Colonial power and African illness*. Stanford, CA: Stanford University Press.

Common Goods

Social scientists call *goods* all the resources that people use in their life. A well-established taxonomy of goods distinguishes them on the basis of two attributes: excludability and rivalry. The first attribute refers to the right of the owner to exclude other individuals from the consumption of the goods. The second means that the consumption of a unit by one individual prevents the consumption of the same unit by another. Private goods, with which basic microeconomics is concerned, are excludable and rivalrous. Goods that lack one or both of these attributes are called common goods by political analysts. This entry discusses the attributes of common goods, a variety of forms of these goods, and ways in which they can be managed.

Three different types of common goods have been envisaged since the 1950s. Economists have drawn attention to public goods, which are nonexcludable and nonrivalrous, such as national defense, traffic lights, air to breathe, and eradication of contagious diseases. More recently toll goods or club goods have been considered, that is, nonrivalrous and excludable goods, such as highways, cable television, shooting or fishing permits, and museums. Finally, environmental scientists and political analysts have envisaged common pool resources (CPRs) or commons, which are nonexcludable but subject to rivalry, such as fisheries, free meadows, hunting game, and groundwater basins. Table 1 illustrates the positions of goods in the taxonomy.

Nonexcludability and/or nonrivalry of common goods pose a series of problems for collective action that have been extensively examined.

Collective Action for Public Goods

Political scientists began to study the role of groups in politics at the beginning of the 20th century. With time, a scholarly tradition became established, and a group theory was proposed that implicitly assumes that whenever all the individuals in a group have the same aim, they will all act to achieve that aim. Mancur Olson, in his book *The Logic of Collective Action*, was the first to argue that this apparently obvious assertion is in general untrue for public goods.

Olson's argument is that the attribute of nonexcludability implies that no individual member of a group has an incentive to contribute voluntarily to the provision of a public good, as each would get his or her part, in any case, once it came into existence. Irrespective of the nature of the group (i.e., a rural community, a town, a union, or a nation), this free riding behavior may spread among individuals, so that a group may be unable to give rise to a

Table 1 Taxonomy of Goods

	Excludable	*Nonexcludable*
Rival	Private goods	CPRs (commons)
Nonrival	Toll goods (club goods)	Public goods

Note: CPRs = common pool resources.

single unit of public good, no matter how important it might be for the group. Consequently, the existence of a public good depends on the capacity of the group to prevent free riding, which means detecting and punishing people who do not cooperate in its provision. This is easiest when the number of group affiliates is small, because they know each other personally, interact frequently, and can regularly monitor their respective actions. Because of this capacity to surmount opportunistic behavior, small groups are privileged and are able to activate and stabilize. Large groups, on the contrary, are anonymous, so that it may be difficult for partners to detect and punish free riding. This is why large groups are difficult to mobilize and may remain latent.

The Tragedy of the Commons

For goods with the rivalry attribute, a unit of CPR taken by one person is taken away from another person. For goods with the nonexcludability attribute, no one can singularly claim property rights. Because CPRs have both these features, they not only have the problem of free riding, shared with nonexcludable public goods, but also are vulnerable to intense exploitation, which may completely erode their value. Garrett Hardin drew attention to this problem in a 1968 article published in *Science*, where he gave the example of an open pasture. Because no individual property right can be claimed on its land, each herdsman will try to keep as many cattle as possible on the pasture. This spontaneous arrangement might last for centuries, as long as wars and diseases keep the population of both men and cattle below the highest level that the meadow can sustain. However, were these traditional evils to be defeated, the population of herdsmen and cattle would grow beyond that threshold, and a new evil, called by Hardin the tragedy of the commons, would emerge.

The catastrophe by no means depends on the carelessness of the herdsmen. Deciding whether to introduce a new unit within the pasture, a perfectly rational herdsman will weigh up expected costs and benefits: Introducing an additional beast gives him the opportunity to add nearly the value of this unit to his income, while the cost of overgrazing produced by that new unit generates only a little erosion of the free grass, a loss that furthermore is distributed among all herdsmen. This reasoning can be extended to any one additional beast and to any herdsman. Therefore, individual rationality pushes each herdsman to increase his herd without limit, progressively reducing the likelihood that any animal will find enough food to survive, and finally, leading all the herdsmen to complete ruin.

Managing Common Goods

In spite of the tragedy of the commons and the problems with collective action, groups of any size have always existed in human societies, delivering common goods to their members in a more or less efficient manner. Of course, this is well known to scholars, who have pondered the various ways in which practical politics has adapted to cope with the problems that theoretical analysis has envisaged.

To begin with, both CPRs and public goods are nonexcludable, which implies exposure to free riding. Then, the theoretical insights considered so far would be corroborated if the strategies of practical politics were to be aimed at increasing the cost of free riding or at giving to common goods some attributes of excludability.

The first strategy amounts to making use of the coercive force of the state—whether or not legitimized by democratic procedures and majority endorsement—in order to deter opportunistic behavior. This strategy is actually largely adopted by all public authorities under all regimes, as the reality of enforced taxation makes clear beyond any doubt. However, the alternative strategy of providing common goods with the attribute of excludability is no less important. Convinced that CPRs generate social evils, Hardin maintained that the history of civilization is that of limiting the "freedom of the commons," introducing property rights by law, and transforming them into private goods, as happens when concessions are granted. Also, public goods can be equipped with elements of excludability. This is accomplished when public authorities introduce fees or tickets that transform public goods into toll goods. Appealing to this strategy, Olson is able to explain how large groups succeed in providing their members with public goods when no mutual control is available to defeat free riding. His explanation is based on the role of selective incentives. These are benefits that the group offers to each potential member as an inducement to join and bear a share of the collective enterprise. Thus, large latent groups have the opportunity to mobilize and gain their public goods as by-products of selective incentives.

An Alternative Way to Cope With CPRs

The strategies so far considered by which groups can provide their affiliates with common goods imply an institutional change, toward either centralization or privatization. However, institutional changes are costly, and in both cases, further analysis of additional costs is needed. This has been clearly pointed out by Elinor Ostrom in her book *Governing the Commons*, devoted to the analysis of CPRs. Her first point is that only public authorities are entitled to introduce property rights on CPRs. That means that creating a market for commons implies the costs that public authorities must bear to establish an agency to define market rules and control their effectiveness and efficiency. On the other hand, if property rights remain in government hands, one must add the cost of detecting defectors as well as the costs of errors, such as mistakenly sanctioning those who cooperate or failing to sanction defectors.

From all this, the conjecture emerges that full privatization or full centralization imposed by external authorities may be a questionable means of coping with the problems with CPRs and that institutional arrangements provided by the commoners themselves may offer better solutions. To examine this hypothesis, Ostrom makes a thoughtful analysis of the decision processes in many empirical cases of CPRs ruled by self-governing communities, such as forest tenures, irrigation systems, or inshore fisheries from various countries in different continents with different levels of economic development. Her scrutiny reveals that people interacting in a CPR can avoid resource erosion through a system of rules contingent on the specific situation, monitored by accountable individuals, and regulated by graduated sanctions. However, to realize such an institutional change, the individuals concerned must know one another through other repeated interaction, so that over time, they can accumulate the social capital necessary to discover shared norms promoting mutual cooperation. These conditions are more likely to be met in the type of small-scale CPRs investigated in Ostrom's book.

It is important to underline the fact that Ostrom does not deny the relevance of the problems that Hardin and Olson have pointed out with common goods. Her aim, more limited and yet extremely important for political analysis, is to show that the assumption that problems with common goods can be solved only by external authorities imposing privatization or centralization, although widespread among policy analysts, is invalid.

Social analysts have scrutinized common goods using contributions of increasing complexity. Hardin's approach exemplifies the environmental scientists' concern for the ultimate tragic consequences of the commons. Olson enhances Hardin's analysis, paying due attention to decision processes and proposing an economic model where individual incentives interact with a given institutional framework. Ostrom considers many aspects concerning the provision and maintenance of common goods. Her analysis strongly suggests that designing the necessary institutional change and considering its costs should not be omitted when modeling decisional processes concerning human attempts to attain collective benefits. Her teachings have been very influential, and subsequent research on common goods has followed her analytic framework, focusing on institutional design and multilevel governance. Thus, it has been possible to extend her method far beyond the small-scale CPRs to research into the international governance of common goods such as the Internet and financial markets and environmental global commons such as Antarctica, the oceans, and the atmosphere.

Paolo Martelli
Università degli Studi di Milano
Milan, Italy

See also Collective Security; Institutional Theory; Social Capital

Further Readings

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Baden, J. A., & Noonan, D. S. (Eds.). (1998). *Managing the commons*. Bloomington: Indiana University Press.

Buck, S. J. (1998). *The global commons: An introduction*. London: Earthscan.
Hardin, G. (1968). Tragedy of the commons. *Science, 62,* 1243–1248.
Hardin, R. (1995). *One for all: The logic of group conflict*. Princeton, NJ: Princeton University Press.
Héritier, A. (Ed.). (2002). *Common goods: Reinventing European and international governance*. Lanham, MD: Rowman & Littlefield.

COMMUNISM

After the fall of the Berlin wall in 1989 and the collapse of the Soviet Union in 1991, it seemed that communism belonged only to the past and was a more appropriate subject of research for historians than for political scientists. This impression is wrong. Communism is directly relevant to political science, not only because communist regimes still exist in Cuba, China, and elsewhere in Asia, but also because an adequate account of communism must include an explanation of why it is attractive or, on the contrary, why it has failed. An understanding of communism must also include an examination of its future as well as its legacy in former communist countries that are now democracies. This entry addresses these topics by analyzing communism as a theory, a type of regime, and a political organization. As a political theory, communism has changed over the years but has always been significant to some philosophers and activists, and it remains an important point of reference in critiques of capitalism and current political debates. In this respect, it is necessary to distinguish between the idea and its political realizations.

Evolution and Actualization of the Idea of Communism

The idea of communism is a very old one that has its roots in the vision of a society based on the absolute equality of human conditions and the elimination of individual enrichment. It is present in the works of many thinkers from Plato through the 18th-century utopians, the Soviet analysts, and some contemporary theoreticians. The terms *communism* and *communist* were first used in France during the 11th century to designate the common practices, interests, and rights of some peasants. During the French Revolution, some authors used the word *communism* in its more modern sense of a general sharing of goods in a regime established by a revolutionary process. Others, for example, Robert Owen, Etienne Cabet, Wilhem Weitling, and Moses Hess, used the term to designate utopian projects of societies based on a new system of exchanges and distribution. But it was around the mid-19th century that the idea of communism became more widespread. The creation of the Communist League by Karl Marx and Friedrich Engels in 1847 and the publication of their *Communist Manifesto* in 1848 marked a turning point in the evolution of the communist idea as embodied in a political structure and metamorphosed into a coherent doctrine that, despite Marx's protests, quickly took on a religious and dogmatic dimension. The word *communism* came to refer to a general project, a political action, and a political ideology presented as a real and powerful science, different from reformism. For Marx, communism meant not only social equality and the end of private property but also the necessity of a revolution based on class struggle directed by an organized proletariat that would completely destroy capitalism and the bourgeoisie, establishing the dictatorship of the proletariat. After this first step, it would be possible to establish a perfected new order with collective ownership, a progressive disappearance of the state, and the dissolution of social classes.

Prior to 1917, there was confusion within the international labor movement concerning the relationship between socialism and communism. After the Russian Revolution of 1917, Lenin gave the name *Communist* to the Bolshevik party, and with the creation of the Third International in 1919, the distinction between the Communist Party and social-democratic and socialist parties was strengthened. Lenin and his comrades set themselves up as Marx's only heirs, delegitimizing the numerous other interpretations of Marxism.

Marxism–Leninism became the main and strongest codification of the doctrine of communism during the 20th century, with the Soviet Union presented as a model. Thus, the Soviet Union can be considered the first incarnation of communism; however, after World War II, other nations, such as Mao Zedong's China, Kim Il-Sung's North Korea, and Fidel Castro's Cuba, created their own versions of communist regimes.

Communism as a Type of Regime

Because of the existence of these countries who claimed their affiliation to communism and Marxism, communism also designates a type of regime. Usually, they show important specificities in terms of how they implemented the communist program and in their respective historical, national, and social environments. However, there were also commonalities; in particular, the establishment of a communist regime was in each case followed by an intense period of violence and terror marked by mass crimes, genocides, trials, deportation, and the creation of detention and concentration camps for all political, cultural, and religious "enemies" as well as bloody purges of opponents inside the party. The prime example was the Soviet Union under Lenin and then under Stalin, until the latter's death in 1953. One of the grimmest examples of communist terror was in Kampuchea (Cambodia), where from 1975 to 1978, the communist Khmer Rouge was responsible for the death of an estimated 1 million to 1.5 million people.

Often, a period of acute repression in a communist regime was followed by a relaxation of the repression; this was the case in Russia and in East and Central Europe after Stalin's death in 1953 and Khrushchev's de-Stalinization in 1956. However, during this period, there remained a restriction of liberties and a general fear based in part on memories of past repression. Other common features are the following:

- the establishment of one-party rule or a hegemonic and preeminent communist party, with some insignificant allies who keep alive the fiction of plurality, as in some of the Eastern and Central European countries after 1945;
- the monopoly of all activities, including political, economic, social, intellectual, and cultural activities, and of the media, who were used for intensive propaganda;
- an emphasis on the importance of ideology for the party and for the whole population;
- the rigorous control and surveillance of society;
- the endeavor to create a "new man," who was ideologically convinced, completely devoted to the party, and willing to sacrifice his life;
- the centralization of all powers, especially in countries with great cultural, ethnic, religious, or geographical diversity; and
- the preeminent affirmation of the leadership, which may lead to an organized cult as, for instance, in the cases of Stalin, Mao, Kim Il Sung, Ho Chi Minh, or Castro.

During the years 1920 to 1930 and after World War II, these features sparked polemics about the nature of communist regimes and their classification: What were they exactly? Were they a new version of classical dictatorship or a variety of authoritarian systems, or did they belong to totalitarianism? The debate about this last notion was always fierce. Supporters of the totalitarian concept insist that it is as valid for communism as for fascism or Nazism. Some, such as Hannah Arendt and Raymond Aron, thought that, after 1956, the Soviet Union, for instance, became a posttotalitarian regime or an authoritarian one that experienced a deep political and economic crisis, resulting in a reduction of its attraction. Other specialists emphasized the permanence of the essential characteristics of totalitarianism from the beginning to the end but recognized that there was alternation between the extreme and the more relaxed forms of totalitarianism. In contrast, other scholars reject the pertinence of the totalitarian concept for communism and criticize the validity of comparing it with fascism. For them, totalitarianism is too wide and schematic a concept for comparison, because it does not take into account the differences between fascism, Nazism, and communism, as well as those between communist countries. These critics also argue that the concept of totalitarianism does not take into account the spaces of autonomy that were operating despite the weight of oppression. In any case, communist regimes also served as a model for dictatorships elsewhere.

Communism as a Political Organization

Communism also refers to an important mutation of political organization. Marx spoke in favor of the importance and the necessity of a communist party but without giving precise definitions or recommendations about its structure. It was Lenin who described more precisely the process of building such a party, in his book *What Is to Be Done?* published in 1902. Lenin is the inventor of the first prototype of an ideological party, composed of a vanguard of professional revolutionaries, with a military and secret service component, publishing a newspaper conceived as a key element of the party's activity, propaganda, and ideological formation. This party was to be ruled such that there was complete submission of the inferior echelons of the party to the superior ones, a strong and centralized authority of the leadership, absolute discipline (whereby the minority had to obey the majority), and, in democratic countries, total control of the members of parliament by the central leadership of the party. This famous invention, called democratic centralism by Lenin, was completely counter to the classical concept of a mass social-democratic party. The party also had to form specialized mass organizations for attracting a variety of people. Through their propaganda, the ideas of communism could be spread more easily, and they could also be used as a channel of recruitment for the party. Trade unions were considered a decisive influence on the working class, but as a transmission agent, they had to be subject to the party's orders. This specific communist party was supposed to be adapted to the conditions of political struggle under tsarist Russia. But after the victorious Soviet revolution, Lenin and Trotsky decided to generalize the model and impose this organization on the new communist parties all over the world. They also created another important rupture with the internationalist working class tradition by inventing a new Communist International, with its powerful center in Moscow. The Communist International was conceived as a world party with one unique goal: to make revolution on a global scale. In reality, with Stalin's leadership from the end of the 1920s, the Communist International had to defend the Soviet Union against the supposed attacks of "imperialism."

The party was completely Stalinized and became monolithic. In other countries, the communist party tried to transform itself into a mass party open to different social categories, for example intellectuals, but with a prominence of working-class members at all levels. In some countries, and especially in China under Mao's authority, priority was given to the peasants. The military had to consist of disciplined members, courageous and always ready to defend the party's positions and, if necessary, to sacrifice their lives, as was the case, for instance, during World War II in the Resistance, during some civil wars (in Spain 1936–1939, in Greece 1946–1949), and in the national liberation movement in Vietnam. The party was attractive because it offered the only possibility for political, social, or symbolic promotion. In the course of time, the ideology was no longer an element of faith but just empty rhetoric, with its sole aim being to legitimize those who were in power. The party was a party-state where the state was only an appearance; the reality of power, leadership, and decisions was under its control. At the international level, the communist parties had been organized from 1919 into the Communist International, the so-called Komintern, which was self-dissolved in 1943. This organization was also Stalinized from the end of the 1920s and had to respect the orders from Moscow. The leadership of the Communist International, directly under Stalin's authority, decided the strategy of each party, controlled the internal life of the party, and selected its leaders. After World War II, the Kominform was created as a successor organization. The Soviet party, with the support of the East European parties, tried to maintain the unity of world communism and the centrality of the Soviet Union by organizing international communist conferences, organizing formal meetings among parties, and providing material (even financial) support. The Soviet party tried to structure this world communist system to achieve its expansion over all continents. But it was confronted with a trend toward autonomy and, sometimes, a defense of national interests (e.g., with the Romanians and the Polish). Some parties decided to oppose Soviet primacy and to split communist unity, such as the Yugoslavs in 1948 and the Chinese and Albanian Communists in the 1960s. In the West, communist Italians after 1956, the French in the 1970s, and then the Spanish soon after tried hard to invent a form of Eurocommunism that emphasized the importance of democracy and personal freedoms and rejected dictatorship. However, they always defended Soviet superiority when confronted with imperialism and capitalism.

Communism as an Ideal

The theoretical, political, and sociological interpretation of communism has always divided the scientific community. Communism has been interpreted as a contemporary form of utopia that would fulfill the classical quest for happiness in a perfect and pacified world and give expression to the "passion of equality" that Alexis de Tocqueville saw as inseparable from democratization. Many thinkers have argued that although communist regimes vigorously repressed religion, communism itself took the form of a secular or political religion. In Europe, communism had more success in Catholic and Orthodox countries and in some Jewish communities than among Protestants. Where Islam was the dominant religion, communism generally failed to take hold; however, it prospered in Buddhist, Taoist, and Confucian countries. Thus, communism could be seen as a secular substitute for traditional religions and its establishment, the result of a transfer of faith from classical religions to this new one.

Communism was embraced by many different sections of the population, from the uneducated to intellectuals, who found in its beliefs a basis for political involvement. Even though it was closely associated with the protection of Soviet interests, communism also may be seen to involve a commitment to internationalism as the foundation of world fraternity. However, communism could also strengthen nationalistic values, as it did in some countries that were fighting for their independence, such as Vietnam or Cambodia. Such ambivalence existed on other topics as well. Communism succeeded as a movement determined to practice violence against dictatorships but also within democracies (as the communist parties did from the 1920s to the 1940s); yet, at the same time, especially in the West, communism presented itself as a peaceful movement allied with Catholic and Protestant associations and independent personalities who were not properly affiliated to a party but were strongly committed to leftist values. Obviously, in theory, communism was a child of the Age of Enlightenment, with its ideals of equality and its aspiration to emancipation and democracy. But in practice, although there were differences between communism and totalitarian movements such as fascism and Nazism, communism translated to dictatorship. In Western countries, communist parties criticized the "bourgeois" and "formal" democracies in the name of a future and perfect democracy. After World War II and their active participation in antifascist resistance or anticolonialist movements, communist parties—the most powerful of which were in Western Europe—gradually accepted democratic principles and made some changes in their own organizations.

But the main question is whether communism represented an archaic political phenomenon or a modern one. Communism, as an affirmation of the values of a classless society, can be seen as archaic because of its rejection of capitalism, the market economy, individualism, and liberal and representative democracy. Communism could also be presented as a specific authoritarian form of economic and social modernization developed by a ruling communist party—a modernization that was a complete disaster everywhere. However, in democratic countries where strong communist parties existed, as in France or Italy, their presence and the protection they gave to the poorest people could be seen as providing a foundation for the modern welfare state.

Communism and Today's World

Communist regimes no longer exist in Russia and in Eastern Europe. Efforts to reform communism in Hungary, Poland, Czechoslovakia, and the Soviet Union under Gorbachev failed. Western European communist parties are in decline or have changed their names and identities. Communism as a whole has lost its momentum. But it is still alive in some countries as a totalitarian regime, as in North Korea, or in the form of coexistence between a market economy and the domination of a single party, as in China. Communism has left an important legacy. Communist domination did leave its mark on welfare in countries that have undergone a transition to democracy, which explains the nostalgia for equality and social protection. In some countries, Communists have maintained a good reputation among the public because the collective memory associates them with the social class struggles of the past and with antifascism. Communist doctrine itself is in a deep crisis, but some authors find an actuality in Marxism. Some elements of the communist culture are still alive in the form of anticapitalism, aspiration to utopian ideals, the search for a radical alternative, contestation of the reality of democracy, and hostility to reformism. In other words, communism is perhaps finished as a political centralized and authoritarian organized movement but remains alive and present as an aspiration.

Marc Lazar
Sciences Po
Paris, France

See also Capitalism; Class, Social; Communist Parties; Communist Systems; Equality; Imperialism; Labor Movement; Maoism; Marxism; Revolution; Socialism; Stalinism; Totalitarianism; Utopianism

Further Readings

Brzezinski, Z. K. (1989). *The grand failure: The birth and the death of communism*. New York: Scribner's.
Courtois, S., Werth, N., Panné, J.-L., Paczkoski, A., Bartosek, K., & Margolni, J.-L. (Eds.). (1999). *The black book of communism: Crimes, terror repression*. Cambridge, MA: Harvard University Press.
Furet, F. (2000). *The passing of an illusion: The idea of communism in the twentieth century*. Chicago: Chicago University Press.
Pipes, R. (2001). *Communism: A brief history*. London: Weidenfield & Nicoleon.
Roberts, D. D. (2007). *The totalitarian experiment in twentieth-century Europe: Understanding the poverty of great politics*. New York: Routledge.
Sandle, M. (2006). *Communism*. Harlow, UK: Pearson Longman.
Service, R. (2007). *Comrades: A world history of communism*. London: Macmillan.
Silvio, P., & Service, R. (Eds.). (2006–2007). *Dizionario del comunismo nel XX secolo* [Dictionary of communism in the twentieth century]. Turin, Italy: Einaudi.
Westoby, A. (1989). *The evolution of communism*. Cambridge, UK: Polity Press.
White, S. (2000). *Communism and its collapse*. London: Routledge.

Communist Parties

Communist parties are radical Leftist political organizations that aim to facilitate either a coercive or peaceful transition of society from capitalism to communism. Such parties initially emerged as factions within many European socialist and social-democratic parties at the end of the 19th century. Communist parties claim to represent the interests and values of the working class. The seizure of power by the Russian Social Democratic Labour Party (Bolsheviks) in a coup d'état in Tsarist Russia in November 1917 was a signal for the emergence of communist parties nearly all over the globe. After World War II, communist parties seized power in Eastern Europe, China, and some other Asian countries. A rapid decline of communist parties worldwide began after Mikhail Gorbachev as General Secretary of the Communist Party of the Soviet Union (CPSU) initiated reforms of the Soviet regime in 1985. This entry discusses their ideology, origins, organization, and worldwide impact.

Ideology

The term *communist* was used for the first time by the utopian socialist Goodwyn Barmby, who introduced the word into the English language as a translation of the French word *communiste* and founded the London Communist Propaganda Society in 1841. An essential distinctive characteristic of communist parties was the importance accorded to ideology. In communist parties, ideology occupies a much more fundamental place than in other parties as a primary concern of the party is to indoctrinate the rank and file with communism or Marxism.

The doctrine of communism refers to the theory of society developed by Karl Marx and Friedrich Engels in the 1840s. In the mid-1840s, Marx and Engels became engaged in politics by joining the League of the Just, a secret society consisting mainly of émigré German artisans. In February 1846, Marx and Engels set up the Communist Correspondence Committee with the aim of reclaiming control of the League of the Just. In June 1847, the League was transformed from a secret society into an open revolutionary organization—the Communist League. The second congress of the Communist League, in December 1847, asked Marx and Engels to draw up a manifesto stating its principles. In February 1848, Marx and Engels published the *Manifesto of the*

Communist Party, where they formulated the basic principles of communism. The major author of *Manifest der Kommunistischen Partei* was Karl Marx; it was originally published in German in London.

A cornerstone of Marxist theory was the economic (often called the *materialistic*) interpretation of history and theory of the social revolution. According to this doctrine, human society is developing through a series of historical stages based on changing modes of production, which are the fundamental determinants of social structures and social relations. As seen by Marx and Engels, the history of human society is the history of class struggles. According to Marxist doctrine, human society has passed through the successive phases of a primitive society, slavery, feudalism, and capitalism and must advance to communism. The class struggle under capitalism is between those who own the means of production—the bourgeoisie—and those who labor for a wage—the working class. Communism means an abolition of private property and the creation of a classless society in which productive wealth is owned collectively.

Origins

Several predecessors of modern communist parties existed, but the major roots of revolutionary movements with a clear communist leaning and a direct involvement of the father of the communist doctrine, that is, Marx, go back to the International Working Men's Association, or the First International. The First International, an organization of workers that represented an alliance of people from diverse groups, such as the French Proudhonists, Blanquists, the English Owenites, the Italian Republicans, proponents of individualistic anarchism, and other socialists of various persuasions, was founded in 1864. Under the influence of Marx, the Social Democratic Workers' Party of Germany, the strongest member of the First International, was founded in 1869. In 1872, the First International split between Marxist and anarchist currents, and the organization was disbanded 4 years later at its 1876 Philadelphia conference.

In July 1889, European socialist and labor parties from 20 countries established the Second International, excluding the powerful anarcho-syndicalist movement. One of the most prominent Russian communist leaders—Vladimir Ilyich Lenin—was a member of the Second International from 1905. The Second International dissolved during World War I, in 1916, as the national parties that composed it did not agree on a unified position against the war.

Marx and Engels argued in the *Manifesto of the Communist Party* that among the immediate goals of Communists were the "formation of the proletariat into a class, the overthrow of the bourgeois supremacy, and the conquest of political power by the proletariat." The ultimate aim of communism was the replacement of the old bourgeoisie by a classless communist society.

Until the collapse of the Second International, Marxists worked within existing social-democratic parties. Marx claimed in the *Manifesto of the Communist Party* that "the Communists do not form a separate party opposed to the other working-class parties." However, in his words, Communists differ from other working-class parties in that they "bring to the front the common interests of the entire proletariat, independently of all nationality," and "they always and everywhere represent the interests of the movement as a whole."

In 1902, Lenin wrote a political pamphlet—*What Is to Be Done?*—where he developed a concept of a "party of a new type." Here, Lenin formulated the basic elements of the concept of a revolutionary party:

a. a centralized and disciplined (cadre) party of professional revolutionaries,

b. a party organized as a bureaucratic or semimilitary hierarchy, and

c. a party able to communicate socialist ideology to the more intellectually developed workers, who could act as the vanguard of the working class and be capable of overthrowing the regime of the bourgeoisie and establishing the dictatorship of the proletariat.

This "Leninist" type of party, as an authoritarian structure, was to become the highest authority of the class-based organization, which would lead all other organizations of the working class and society as a whole. For many decades, the organization, strategy, and tactics of the Leninist party were interpreted as providing the impetus for the establishment of the totalitarian regime in the Soviet Union.

In 1990, however, the American socialist Hal Draper published an article "The Myth of Lenin's 'Concept of The Party' or What They Did to *What Is to Be Done?*" where he argued that Lenin's concept of the working-class party was highly misinterpreted. First of all, Lenin did not speak about intellectuals as the carriers of socialist ideas in the political organizations of the working class. Second, Lenin's demand of centralism did not mean a "democratic centralism"—that is, a semimilitary hierarchy within the party—but rather the aim to establish an organized all-Russian party for the first time. The aim of having a class-based party was just a reaction to the fact that throughout history, the socialist movement was mostly organized in sects. Third, Lenin did not propose a revolutionary sect, but instead, he aimed "to organize the revolutionary current as a political centre of some sort inside the mass party." John P. Plamenatz (1954), a political philosopher, argued long before Draper that in 1902, Lenin's ideas of a revolutionary party were not "undemocratic, but merely advice adapted to the needs of a revolutionary party active in Russia in the first decade of the twentieth century" (p. 225), which, however, took an antidemocratic turn in "what happened after the Bolshevik Revolution."

In 1918, the Russian Social Democratic Labour Party (Bolsheviks) was renamed the Russian Communist Party (Bolsheviks), and from 1920 until 1991, it was the only legal political party in the Soviet Union. In 1925, the name of the party was changed once again, and it became the All-Union Communist Party (Bolsheviks) and eventually was transformed into the CPSU in 1952. After the attempted *coup d'état* in August 1991, the CPSU was dissolved by Russia's President Boris Yeltsin.

Organization

As the first communist parties were splinter groups of existing social-democratic and/or socialist parties, initially, they inherited the traditional organization of mass parties. After the October Revolution of 1917, the Russian Communist Party gradually centralized political power and developed a strict cadre policy. A high degree of centralization is a distinctive characteristic of communist parties. All mass-based parties tend to be centralized, but within communist parties, it was significantly more important than in other mass-based parties.

All political factions and dissent were forbidden within the Russian Communist Party in 1921, and afterward, all party members had to adhere to the decisions that had been made by the central bodies. The internal party organization was based on the principle of democratic centralism, where "democracy" meant the freedom of party members to discuss matters of policy and to pass decisions by majority vote, and "centralism" was understood as a duty of party members to uphold decisions taken by the majority. In theory, democratic centralism also included elections to all party institutions from the bottom to the top, accountability of party structures and leaders to the rank and file, strict party discipline and mandatory decisions of the upper structures for the lower structures, and personal responsibility of party members for the assignments given to them by the party. In reality, democratic procedures were an empty formality as leadership at all levels of the party organization prohibited criticism and dissent in decisions and actions.

The Russian Communist Party also introduced a *nomenklatura* system through which party elites were appointed to key positions in the administrative, economic, military, cultural, and educational bureaucracies as well as throughout the party's own hierarchy. Lenin himself formulated the first criteria for *nomenklatura* appointments as follows: reliability, political loyalty, and administrative ability. Joseph Stalin finished the building of the party's patronage system and used it to distribute his confidants throughout the party bureaucracy. Under Stalin's direction in 1922, the party created departments of the Central Committee and other organs at lower levels that were responsible for the appointment of party officials. After Leonid Brezhnev's accession to power in October 1964, the party considerably expanded its appointment authority. At the all-union level, the Party Building and Cadre Work Department supervised *nomenklatura* appointments. This department maintained records on party members throughout the former USSR, made appointments to positions on the all-union level, and approved *nomenklatura* appointments to the lower levels of the hierarchy. The head of this department was sometimes a member of the party Secretariat and was often a protégé of the General Secretary of the CPSU.

The communist parties developed a new structural organization. While other mass-based parties focused their organization and drew their support from a particular geographical or territorial area, the communist groups formed their cells in the workplace. Party members were employed in the same enterprise or professional institution, above which were territorial district, city, regional, and national party committees. The workplace cell organization proved to be effective. These workplace cells were more numerous than territorial local branches in other mass-based parties' sections. Also, this system led each cell to concern itself with problems of a corporate and professional nature rather than with more political issues. Last but not least, this organizational system allowed party leaders to have extensive authority over primary party cells. At the end of the 1980s, the CPSU had around 20 million members organized into about 400,000 primary party cells.

At the top was the party congress, which met only every few years, where strictly selected delegates elected the members of the republican and the all-union party central committees. The Politburo was the highest political and executive institution of the CPSU. It was created on the eve of the Bolshevik Revolution in 1917, but it became fully functional only in 1919 and soon overshadowed the Central Committee in power. Nominally, the Politburo was elected by the Central Committee to direct the party's policy between the Central Committee's plenary sessions. In fact, the Politburo was the most important decision-making institution in the Communist Party and has commonly been seen as a rough equivalent to the cabinet in Western political systems. The size of the Politburo of the CPSU varied, and from 1952 until 1966, it was called the Presidium of the Central Committee.

In addition, the Communist Party was surrounded by various satellite mass organizations. In the Soviet Union, among the most important were the Communist Youth Union (Komsomol), from which nearly 75% of party members were recruited, and the All-Union Central Council of Trade Unions, with more than 100 million members. The Communist Party controlled all facets of economic, political, military, and cultural life, and Article 6 of the 1977 Soviet Constitution stated that "the leading and guiding force of the Soviet society and the nucleus of its political system, of all state organizations and public organizations, is the Communist Party of the Soviet Union."

Communist Parties Worldwide

After the establishment of the Communist International (Comintern) in 1919 by representatives of 34 parties, most emerging communist parties worldwide were patterned along the model of the organization of the Russian Communists. In 1920, the Second Congress of the Communist International formally adopted the *Twenty-One Conditions*—officially the *Conditions of Admission to the Communist International*—as prerequisites for any political group wanting to become affiliated to the Communist International. These conditions encouraged the buildup of communist organizations along democratic centralist lines.

Many European Socialist parties went through splits based on whether they adhered to the Communist International. The French Section of the Workers International (SFIO) broke away with the 1920 Tours Congress, leading to the creation of the new French Communist Party. The Communist Party of Spain was created in 1920 on the basis of the Federation of Socialist Youth; the Communist Party of Italy was created in 1921 following a split in the Italian Socialist Party; the Belgian Communist Party emerged in September 1921 following a split from the Belgian Workers Party, and so on. Membership in the Communist International in itself formalized the split between the Communists and moderate social democracy.

In the interwar period, the Communist Party of Germany was not only a major party in Germany but was also the strongest communist party in Europe outside of the Soviet Union. In the four Reichstag elections of 1930–1932, the Communist Party polled between 5 and 6 million votes. The party was then suppressed by the Nazi regime. A considerable number of senior German communist leaders in exile were also executed during Stalin's Great Purge of 1937–1938, and more than a thousand were handed over to Nazi Germany by Stalin's regime. The Soviet-dominated communist parties and the Comintern in the interwar period were the main instruments to export the Bolsheviks Revolution. By the end of the 1920s, the revolutionary upsurge in Europe was over, and communist parties were established in most countries, but in most cases, the Communists did not play a leading role in the labor movement.

After World War II, the Soviet victory over fascism contributed to a surge of popularity of communist parties, even in Western Europe, which were originally established in 1918 to 1923. Communist parties were stronger in the less developed countries of Southern Europe, particularly in Greece and Italy but also in France. However, their influence significantly decreased from the 1980s onward.

The Italian Communist Party was established in 1921 through a split within the Italian Socialist Party. It was outlawed by the Mussolini regime but eventually reemerged as the strongest party of the Left in Italy, and at its peak, it was the largest communist party in Western Europe, with a membership averaging about 1.5 million from the 1940s until the early 1980s. The party was organized like the CPSU, with local sections forming federations, which in turn were grouped into regional committees. Since the late 1940s, the Italian Communist Party has held power in many municipalities, and since the 1970s, it has shared control of major urban centers in the country with the Socialist party. However, much of postfascist politics in Italy was organized around the perceived need to keep the Italian Communist Party out of central government.

The French Communist Party was founded in 1920 as an antisystem, leftist, and secular political organization representing French workers and became especially strong in the industrial areas of northern and eastern France and the suburbs of Paris, obtaining a high of almost 29% of the votes in the 1946 elections. Although the French Communist Party has never won a majority of voters, by the late 1970s, it was the largest of all French parties, with a membership of some 700,000, and a few Communists served in the leftist cabinets in 1981 to 1984 and in 1997. Beginning in the 1970s, some currents in the French Communist Party joined with those of Italy and Spain to advocate a more liberal and pluralistic form of communism, that is, Eurocommunism.

Greek Communists, similar to those in France and Italy, built their leading role on the claim of anti-Nazi resistance during World War II. In 1944, Greek Communists participated in the national unity government, holding a number of important portfolios, including the position of Minister of Finance. After the Greek Civil War in 1946 to 1949, the Communist Party was banned until the restoration of parliamentary democracy in 1974. Two other Southern European communist parties—the Spanish and the Portuguese—were politically strong and at their peak after the transition to democracy in the 1970s. They captured about 15% and 20%, respectively, of the total vote in national elections.

In Western and Northern Europe, with the exception of Finland, communist parties had only a marginal political influence. In Finland, Communists created the Finnish People's Democratic League, which received around 20% of the vote on a regular basis till the end of the 1970s.

Eastern Europe, with the exception of Yugoslavia, became a bastion of Soviet-style communism, as communist parties, relying in most cases on Soviet military presence, were able to win and consolidate power in 1945 to 1948. Nonetheless, all the communist parties of Eastern Europe, with the exception of Albania (the party was founded in 1941), had their origins in the period 1891 to 1921. Communist-led governments after 1945 were formed in East Germany, Hungary, Czechoslovakia, Bulgaria, Romania, and Poland. In some cases, communist parties cemented their power through a merger with other socialist or social-democratic parties, as in the case of the Socialist Unity Party of Germany, the Hungarian Working People's Party, the Romanian Workers' Party, and the Polish United Workers' Party. Soon after coming to power, ruling communist parties abolished free elections and other political freedoms. Attempts to reform the existing regime in the Communist-ruled countries were suppressed directly by the intervention of Soviet-led military forces—as in the German Democratic Republic in 1953, Hungary in 1956, and Czechoslovakia in 1968—justified by the Brezhnev Doctrine of the limited sovereignty of socialist countries or by the repression of the national communist government, as in Poland in 1981 to 1983.

However, the CPSU was not able to maintain unity and control of the international communist movement and communist regimes. As early as 1948, the Yugoslav Communists removed themselves from Soviet subordination. Yugoslavia and Albania were liberated by communist partisans (led by Josip Broz Tito) and without Soviet military aid. This was the main reason to stay largely independent of the CPSU hegemony and to advocate national communism as an independent road to socialism. Albania remained allied with the Soviet Union from 1945 to 1960, but the Party of Labour of Albania was Stalinist even after the CPSU had turned away from Stalinism in the late 1950s. After 1960, it broke with the former USSR and joined ranks with Maoist China.

The supremacy of the CPSU was most challenged by the Chinese Communist Party (CCP). The party was founded in 1921. In 1949, the Chinese Communists were able to win power on their own and established the People's Republic. Since 1949, the CCP, initially led by Mao Zedong, has been the country's only legal party. Organizationally, the CCP followed the model of the CPSU, having at the top a Politburo and Central Committee, together with party committees at the provincial, district, county, and municipal levels. Divergence between the Chinese and Soviet parties gradually emerged from the mid-1950s through ideological disputes and Khrushchev's policy of de-Stalinization; later, it led to a Cold War–like conflict between the two communist nations. At the end, the CCP tried to win over to its side communist parties throughout the world. This was most successful in Southeast Asia and in Albania.

In some Latin American countries, notably Chile, El Salvador, Brazil, Mexico, and Venezuela, mass communist parties exercised considerable political influence but never held power as a single party. They supported, however, various leftist coalitions and sometimes guerrilla activities. The only ruling communist party in Central America is the Cuban Communist Party, which came to power in 1959 through a revolutionary movement led by Fidel Castro. The Cuban Communists were heavily dependent on the former USSR for financial support during the 1970s and 1980s. The Cuban communist government was able to survive the collapse of communism in Eastern Europe after 1989, and Fidel Castro ranked among the world's longest-serving heads of state.

In the 1980s, more than one fourth of the world's population lived under communist rule. Two of the world's most populous nations, China and the former USSR, had communist governments. Communist parties held power in Asia (Afghanistan, Cambodia, Laos, Mongolia, North Korea, and Vietnam), many countries of Eastern Europe (Albania, Bulgaria, Czechoslovakia, East Germany, Hungary, Poland, Romania, and former Yugoslavia), Africa (Ethiopia), and Central America (Cuba).

The communist parties, despite the same or similar principles of organization and similar ideological beliefs, have never been a fully uniform or monolithic phenomenon. It is possible to point out the following distinct phases of development of communist parties worldwide:

a. a period of establishment of these parties as autonomous political organizations in the second and third decades of the 20th century;
b. Bolshevikization and Stalinization of the communist parties in the 1920s and 1930s;
c. a challenge to Soviet hegemony in the communist movement of the late 1940s and 1950s through the ideological split between Tito and Stalin and the rise of Maoism;
d. the de-Stalinization of the CPSU and most Eastern European communist parties in the late 1950s and early 1960s;
e. the rise of Eurocommunism, or "socialism with a human face," primarily in Southern Europe since the 1970s; and
f. perestroika in the Soviet Union and the end of orthodox communism in Eastern Europe in the late 1980s and early 1990s.

Altogether, the 20th century could be called the "age of shine and decline" of communist parties worldwide.

Communist Parties After 1989

With the collapse of communism in the Soviet Union and Eastern Europe, communist parties were severely weakened throughout the world.

CPSU membership declined from about 19.5 million in 1988 to 15 million in 1991, and after the attempted coup d'état in August 1991, it was finally dissolved. Responding to democratic upheavals in the former USSR, European and other communist parties redefined their ideological positions moving in a direction of social democracy. Some of them changed their names, and some disappeared from the political scene altogether.

However, at the end of the 20th century, the transformation of the communist parties was not a linear process. Unlike the communist organizations of Eastern Europe, the CCP was able to stem the democratic protest in the late 1980s and introduced economic reforms from above, which led to an impressive modernization of the country. The CCP is the largest communist party in the world, with an estimated membership of more than 66 million at the beginning of the 21st century.

In some Eastern European countries such as Bulgaria, Lithuania, Poland, and Romania, communist organizations were transformed into social-democratic parties, and these successor parties have achieved significant electoral success. In many post-Soviet countries, except the Baltic States, communist parties survived the demise of the Soviet Union as they were reorganized along national lines. The Communist Party of the Russian Federation was (re)founded in 1993. In general, these parties preserved certain leftist ideological beliefs and regained reduced political influence emphasizing nationalism instead of communist ideology (in Moldova, Ukraine, and Russia).

In recent years, in a number of countries such as Nepal, Cyprus, Brazil, India, Venezuela, South Africa, and Sri Lanka, Communists have enjoyed a certain level of electoral success and shared power in multiparty national and local governments. In five countries—China, Laos, Vietnam, Cuba, and North Korea—communist parties remain in power. However, three countries—that is, China, Vietnam, and Laos—have moved toward market economies but without major privatization of the state sector.

Algis Krupavičius
Kaunas University of Technology
Kaunas, Lithuania

See also Communism; Maoism; Marx, Karl; Marxism; Social Democracy

Further Readings

Claudin, F. (1975). *The communist movement: From Comintern to Cominform*. New York: Monthly Review Press.

Draper, H. (2000). The myth of Lenin's "concept of the party" or what they did to *What Is to Be Done?* Retrieved October 15, 2010, from http://www.marxists.org/archive/draper/1990/myth/myth.htm

Guiat, C. (2003). *The French and Italian Communist parties: Comrades and culture*. Portland, OR: Frank Cass.

Laird, R. D. (1986). *The Politburo: Demographic trends, Gorbachev and the future*. Boulder, CO: Westview Press.

Lowenhardt, J., Ozinga, J. R., & van Ree, E. (1992). *The rise and fall of the Soviet Politburo*. London: UCL Press.

Plamenatz, J. P. (1954). *German Marxism and Russian communism*. London: Longmans.

Communist Systems

Communist systems were historically shaped in the 20th century and based on oppressive centralized political power aiming at control over a usually nationalized economy, culture, and society. Such systems appeal to a communist ideology according to which a fully communist system ensures the liquidation of exploitation and equality in a classless society. Because this ideology remained an unfulfilled utopia, the term *communist systems* refers to existing systems based on the characteristics mentioned above: oppressiveness and striving for control over a nationalized reality. According to communist ideology, such systems are supposed to be only transitional forms on the path to full communism, while in reality they are the only form of communism in existence. Although this entry discusses communism considered as a system rather than an ideology, it should be remembered that the communist ideology caused the deaths of tens of millions of victims of communist repression. This ideology was responsible for economic waste and the stifling of human energy that led to the collapse of the system in Europe in 1989 and the fall of the Soviet Union, leaving only a few communist countries on the map of the world.

Communism was a peculiar system in which formal solutions generated their opposites, as many researchers noted (e.g., Alain Besançon, Elemer Hankiss, Robert Sharlet, Jadwiga Staniszkis). Centralization gave rise to spontaneous decentralization, formal regulation resulted in informality, and ideologically based unity brought a diversity that could not be coped with. This is an example of the dynamic of this system. What created communism as a system was an ability to generate and absorb these contradictions in a single whole. A system driven by contradictions was integrated by the relations between particular poles of opposition, with the result that both opposing poles were in need of each other and were strengthened. This created communism as a system and allowed it to reproduce itself. At the same time, however, these contradictions were also a factor for change and adaptation of the system. The informal economy and its relation to the official economy can serve as an example here. The informal economy made it possible to function and somehow also changed the system but, by consuming resources, also contributed to its erosion. These contradictions of the official form of communism were the result of its inability to adapt to social needs and aspirations, and as a result, the contradictions were tolerated by the official system—a mechanism of adaptation extending the life of the system and enabling reproductions.

This entry presents three basic areas of function of communist systems: politics, the economy, and society. Making a precise separation of the three areas is difficult because the specificity of communist systems lay in the powerful tangle of political economic and social structures. At the same time, there was a hierarchy: Politics was the prime cause overriding the other areas.

From Dominance by Politics to a Lack of Politics

In the institutional sense, communism was a system based on the domination of politics over other spheres of life, the organization of life being based on principles of centralism and the elimination of political and social pluralism. One consequence of the aim to eliminate independent structures was an economy based on state ownership, but the economy is just one example of the attempts to subordinate social life to the political authorities. The system sought its justification in an ideology that derived communism from criticism of the structural contradictions of capitalism and promised the creation of a society free from contradictions. Communism in reality has sometimes been seen as a system fulfilling the requirements of the totalitarian model, something rather doubtful in view of the many internal contradictions mentioned above. The totalitarian model then can be seen as a certain type of communist system.

An important characteristic of communism as it actually existed in the political dimension was the paradoxical coexistence of revolutionary mechanisms in the power system and the façade of institutions of parliamentary democracy remaining from the previous system. Although the communist party was the dominant actor, the existence of a government and parliament provided a façade of parliamentary democracy. Its main function was legitimation, although it was also the meeting point and place of negotiation for various group interests, above all at the local level. The communist system, based on the assumption of a unity of interests in society—because after the elimination of antagonistic classes, the basis for opposition of interests was supposed to disappear—had, however, to deal with their real differentiation. These interests were various but in a significant way arose from the segmentation introduced by the system itself, in various branches and sectors throughout the institutional whole. Communism was unable to suspend the working of the laws of economics, which is why real if not overt negotiations and peculiar bargaining appeared in the economy. It was also not able to suspend the generation of interests by all organizational structures—interests in a large part autonomous with respect to central decrees. In an informal way, they had therefore to penetrate the formal structures. This created fertile ground for phenomena related to corruption.

The necessity of taking into account the differentiation of interests in the political sphere in fact marks one of the institutional contradictions of the communist regime. On the one hand, there was the principle of centralism, while on the other, to maintain its ability to reproduce and adapt itself, it was necessary to incorporate into its structure the differentiated interests, at least to an extent that enabled it to continue functioning. The coexistence of those incoherent elements constituted both an opportunity for reform and a limitation on the reformability of the institutional system of communism. The opportunity was created by the incompletely defined character of the boundaries, and the limit was the fundamental contradiction: The more bottom-up representation of interests was strengthened, the more central control was weakened. This mechanism was clearly visible in the case of unsuccessful economic reform.

The aim of total political control not only resulted in its opposite but also led to the disappearance of any meaningful politics. A system in which everything is controlled by politics in fact does not leave room for any politics at all. A lack of real politics, understood as negotiation of pluralistic interests, seems obvious in the case of communism—in which, paradoxically, everything was meant to be politics.

From a Centrally Planned Economy to Disorganization

The economy was one of the spheres in which the communist system was supposed to fulfill ideological aims. In practice, this was a matter of subordination of economic choices to political criteria. This mechanism was based on institutional structures joining economic organisms with the administration of the state, which moreover was the chief owner of the economy. The basis for economic choices and decisions was a system of economic plans whose justification was the overcoming of the irrationality and contradictions of the capitalist economy and the ensuring of social needs by an economy of real socialism. This system created institutional frameworks for the so-called socialist industrialization, which according to the Polish sociologist Witold Morawski was characterized by preeminence of the means of production, central planning, hierarchic administrative control, strong engagement of the party in the economy, and a lack of independent economic institutions.

The striving for political control was manifested in management and planning carried out centrally. The centralist way of planning and the command-distribution system of management was justified strictly by economic necessity; sometimes, it was argued that it was necessary in a situation of limited resources. In practice, it turned out that linking an economy of scarcity to centralization produced the opposite of what was intended. Centralization failed as a mechanism of control precisely when resources were scarce, and it transformed itself into its own opposite, into spontaneous decentralization. Sociologists such as Lena Kolarska-Bobińska have shown that in such a situation of an economic scarcity, real influence was gained by suppliers of resources at the cost of the political authorities, and this is a way of decentralizing without a precise awareness. As a result, the system of management of the economy, instead of being a system of its hierarchical control became instead the space for various games in which organizations and various interest groups within the institutions of management participated. However, that spontaneous decentralization did not damage the essence of the system because it still had at least two mechanisms of control. The first was control by personnel policy, including the *nomenklatura* mechanism, through which the party controlled its personnel. Perhaps, the economy was uncontrollable, but that did not mean that some other source of control gained power. The economy drifted rather than moving in another ideologically unacceptable direction. The economy was further limited by the fact that due to communism's definition of the preeminence of the branch producing the means of production, suppliers of precisely that branch gained more influence and indirectly reinforced the strategic aim of control over the economy. That control at best became more static than dynamic; the economy did not achieve the set goals.

A second source of limitation on the centralizing mechanisms was the cycle of reform. To satisfy the needs of society, the communist economy had to realize economic aims, not merely political ones. The implementation of communism showed that these aims remained in contradiction with one another, and in the last instance, political aims took precedence. The typical cycle of unrealized reform resulted in low effectiveness of management leading to the danger of social upheaval, which was quite frequent in some communist countries (e.g., Poland). In response to this threat to its power, the current government instituted economic reform by loosening its regulation of the economy; however, this decreased control of the economy, which, in turn, heightened the government's fear of losing power, and therefore, the government retreated from this economic policy. In this way, the same political factor (fear of loss of power) led both to the initiation of reform and to withdrawal of the same. This does not mean, however, that the reforms had no effect. Each led to the discovery of successive limits to the reformability of the system. At different points of time during communist rule, various provisions were considered as solutions to existing problems, such as change of cabinets, managerial reform (more independence for directors of firms), and a role for mechanisms of self-government. None of these reforms on its own was introduced systematically, but together, they had a cumulative effect. That accumulation of experience of unsuccessful attempts at reform led in the end to the discovery that it was impossible to reform communist economic policy.

The next characteristic of the communist model of management was emphasis on formal procedures and mechanisms that also generated their own contradictions in the form of the role of informal mechanisms. The inability to satisfy the needs of the population generated various informal mechanisms that can be regarded as types of supply structures. The informal economy is compatible with different systems, but under communism, it fulfilled a specific function. This is because its essence was not avoidance of registration for reasons such as taxation, as happens in market economies. In the nonmarket economy, it served as a mechanism compensating for the deficit of market goods. From this point of view, it can be said that it constituted a necessary supplement to the economy of "first circulation" to a certain degree tolerated by the official system. Therefore, it is possible in fact to talk about a powerful entanglement of formal and informal structures in which the existence of the latter was an important element of the economy.

However, not every deviation from the economic aims of enterprises was a result of informality. Some were direct results of ideological assumptions according to which the workplace was supposed to be a forum for achieving not only pragmatic but also social aims: the socialization of employees to function in conditions laid down and preferred by the system. Creation of the real communist economy was at the same time an attempt to create a class of socialist employees. For this reason, there were various forms of competition, ideological schooling, and methods for satisfying the needs of employees in their role as consumers, not through the market but through the workplace. The whole system of the communist welfare state was based to a large extent on the distribution of various goods through the workplace. The provision of health care, access to cultural events, and education were often tied to the workplace. This social function of the workplace—constituting an element of the ideology—was at the same time functional in addressing the permanent state of economy of scarcity. Distribution through the workplace of consumer goods in short supply was in fact a supply mechanism of the economy of scarcity. Simultaneously, it found its justification in the ideology of the system.

Society: From a Myth of Unity to Real Contradictions

The model of communist society in its actual form never achieved full unity. This section first outlines the mechanisms creating inequality into a system appealing to the idea of equality and then discusses the rejection of the legitimization of the system that according to its ideology rested on mass support.

The real communist principle of "from each according to his usefulness to the authorities" infringed the principle of "from each according to his work," which was an element of the ideological basis of communism. The two therefore were in contradiction and, what is more, infringement of either of them brought about a crisis of legitimization with respect to the basic principles of the identity of the system. The mechanisms of the deviation from the principle of "from each according to his work (ability, contribution)" were largely caused by politics that by a system of priorities favored defined institutional segments of the system. In the economic sphere, it was the branch of production of the means of production, and above all heavy industry, and those segments of machine building that were useful for military production. This was not just a matter of ideologically motivated segmentation. In the case of the real economy of scarcity, branches or indeed firms producing scarce consumer goods took on particular importance.

This whole mechanism demolished in an obvious way the official vision of social stratification. It led to the decomposition of status factors described in the literature by shaking the tie between contribution of work measured in terms of—for example—qualifications and prestige or earnings. Instead, in the definition of the level of income from work, or more broadly the economic position of the individual, an essential role was played by the branch of industry or indeed the type of organization in which one worked. Systemic sources of social inequality in communism were supposed, in theory, to ensure the political stability of the system but, in practice, led to further erosion of its legitimization. This is because it was not possible to justify any of the legitimizing principles—by appeal either to the slogan "from each according to his work" or to the slogan "from each according to his needs." This is another example showing how the adaptive mechanisms generated by systemic incoherence sustained its further reproduction in the short term while, in the longer term, contributing to its collapse.

All this revealed that mass legitimization (the conviction that subordination to the system is morally justified) was a myth to be replaced by multiple processes of adaptation. Various authors have shown that adaptation rather than legitimization was the factor maintaining stability, albeit within a limited scope and for a limited time. It lasted for some time, however, and those factors of accommodating stabilization were varied and evolving. There were attempts to substitute economic achievements for political legitimization, with a relative quasi legitimization or various forms of pragmatic adaptation to the system.

Without doubt, the establishment of the Solidarity trade union in Poland in 1980 dealt a vital blow to the legitimizing myths of communism. This was because for the first time, workers questioned the myth of unity of interests in an open and in fact institutionalized manner. They demanded neither joint management nor participation in decisions, as had occurred after 1956 on the wave of popularity of self-management. Rather, the demand was for trade unions as separate organizations that would serve their interests as employees, interests that they did not see as identical to those of their employers. This was the beginning of the end of the system based on the myth of unity of those interests because the state under real communism was supposed to be precisely the representative of employees' interests since that was why private ownership of the means of production was abolished. In this way, the legitimizing myth of communism came to an end. Before that happened, however, communist systems were torn by waves of outbreaks of social dissatisfaction (e.g., during 1956 in Hungary and in 1970, 1976, and 1980 in Poland) prompted by frustration with an inefficient economy that failed to meet workers' interests and economic needs. By degrees, however, these revolts took on a political character and articulated a demand for institutional change, one example of which was the creation of trade unions independent of the authorities, as noted previously.

Variations in Communist Systems

Communist systems were not static; they underwent change over time. This dynamic was one of the two fundamental sources of their variety. The second was differentiation resulting from various traditions and social and cultural conditions. The first communist state constituting a sort of template of the system was the Soviet Union. (Formally, it was established in 1922; however, Soviet Russia came into existence in 1917 and was not dissolved until 1991.) The government of Joseph Stalin (1878–1953), who led the Communist Party in the USSR from 1922 until his death, can be seen as the formative period of the institutions of the communist system. This period coincided with the building of the military position and the significance of the communist system that controlled many countries of Central and Eastern Europe after World War II (1939–1945), realizing the imperial mission of the Soviet Union in the face of the passivity of the West with respect to the expansion of the USSR, recently an ally in the war against Nazi Germany. The building of this position was achieved at the cost of millions of victims of a whole system of repression comprising police terror, administration of justice entirely under political control, and a system of labor camps. Estimates of the number of victims of the entire communist system fall, according to Jean-Louis Margolin and

Nicolas Werth, in the range of between 65 million and 93 million (Stephen Courtois et al., 1999). After the death of Stalin, the communist system underwent gradual liberalization, with a somewhat looser political control over different spheres of life and a weakening of the system of mass repression against individuals and groups seen as enemies of the system. Even though the wave of terror and repression of Stalinist times did not return, nevertheless, the systems never achieved complete liberalization because their full realization would have simply meant questioning their very essence.

Apart from changes over time, it is also worth noting local variations in the communist system. Communism was introduced by force after World War II in many countries of Eastern and Central Europe and was brought to an end by the peaceful revolutions of 1989. Since 1949, a communist system has prevailed in China, which today has a population of more than 1.3 billion; communist systems also exist today in North Korea and Cuba. There is therefore a group of communist countries with varied traditions and social and ethnic structures. Communist systems within Europe were also quite varied. For example, the Romanian variant, which maintained a certain independence from the Soviet Union, was much more repressive than the Hungarian or Polish versions, while the Yugoslavian variant in turn strongly accented self-management (under the control of the Communist Party, however). Taking the Asian variants into account only deepens the differentiation. The Chinese model links a powerful, almost fully market economy with political control by the Communist Party; in contrast, the North Korean model approaches the totalitarian model in the scale of its control and repression. In spite of the differences among communist systems, the limits to change were defined, and attempts to cross them met with refusal. They were emphatically manifested in the military interventions by the Soviet army in Hungary in 1956 and in Czechoslovakia in 1968 (when the Soviet army was accompanied by those of a group of countries from the Warsaw Pact) and also the introduction of martial law by the Polish Communists in 1981. In each case, these actions constituted a brutal reaction to attempts to liberalize the communist systems.

Collapse of Communist Systems

The sources of the downfall of the communist system lie above all in its inability to accommodate itself to social needs and aspirations. This brought about many manifestations of dissatisfaction and the establishment of movements of opposition to communism in many countries (e.g., Sajudis in Lithuania, Charter 77 in Czechoslovakia, and the Workers' Defence Committee in Poland). The catalyst for these changes was without doubt the disintegration of the Soviet Union, being an unintended effect of reforms begun by Gorbachev, leader of the Soviet Communist Party in the last period of existence of the USSR. Additional factors in the erosion of communist systems were the opposition to communism and the diplomatic efforts of Pope John Paul II. It is a unique phenomenon in the history of civilization that basically, apart from certain exceptions, the collapse of the Soviet Union and the communist systems in the European countries took place peacefully, mainly as a result of negotiation by opposition elites with the communist authorities.

Communist systems, then, had internal sources of dynamics and change. The dynamic was neither ideologically based, nor one consisting exclusively of constant protest under which communism eventually folded and collapsed. It was also a matter of a dynamic resulting from contradictions generated by the system itself that turned out in the end to be insurmountable and that led to its collapse.

Not all the cracks in the systems arose from contradictions generated by the dynamic of communism: Some of them were the result of the persistence of certain systemic remnants from the previous regime (in certain countries, e.g., in Poland, a large role was played by private ownership of agricultural land and by the Roman Catholic Church).

In this way, then, the processes of adaptation that made continuation of the system possible, in the end, also constituted the seed of its collapse, as they led to erosion of resources that finally rendered the system incapable of production. They were, therefore, the source of the persistence of the communist system, its modification, and finally its collapse.

Andrzej Rychard
Graduate School for Social Research
Warsaw, Poland

See also Authoritarian Regimes; Communist Parties; Socialist Systems; Totalitarian Regimes; Totalitarianism

Further Readings

Courtois, S., Werth, N., Panne, J.-L., Paczkowski, A., Bartosek, K., & Margonin, J.-L. (1999). *The black book of communism: Crimes, terror, repression.* Cambridge, MA: Harvard University Press.

Furet, F. (1999). *The passing of an illusion: The idea of communism in the twentieth century*. Chicago: University of Chicago Press.

Holmes, L. (1997). *Post-communism: An introduction.* London: Polity Press.

Kamiński, A. Z. (1992). *An institutional theory of communist regimes: Design, function, and breakdown.* San Francisco: ICS Press.

Staniszkis, J. (1992). *The ontology of socialism.* Oxford, UK: Clarendon Press.

Tucker, R. C. (Ed.). (1977). *Stalinism: Essays in historical interpretation.* New York: W. W. Norton.

Communitarianism

Communitarianism is a social philosophy that builds on the assumption that the good should be defined socially. This core assumption is in sharp contrast with liberalism, which assumes that each person ought to determine the good individually. Communitarianism stresses that individuals are socially "embedded" rather than free agents, that people have social responsibilities to each other, to their communities, and to the common good.

Although communitarianism is a small philosophical school, it has a measure of influence on public dialogues and politics, especially as an antidote to the laissez-faire conservatism championed by Margaret Thatcher in Britain and Ronald Reagan in the United States. President Barack Obama gave voice to communitarian ideas and ideals in his book *The Audacity of Hope* and repeatedly during the 2008 presidential election campaign, calling on his fellow citizens to "ground our politics in the notion of a common good" (2006, p. 9), for an "age of responsibility" and to forgo identity politics in favor of building community-wide unity. At the same time, it should be noted that many in the West consider *communitarian* an awkward term that evokes misleading associations. Hence, Obama, as well as other public leaders who have embraced communitarian themes, such as Tony Blair in the United Kingdom and Bill Clinton in the United States, avoid the term itself.

Branches of Communitarianism

Academic (Philosophical) Communitarianism

In the 1980s, communitarian thinking was largely associated with the works of political philosophers such as Charles Taylor, Michael Sandel, and Michael Walzer. Others sometimes associated with this group include Alasdair MacIntyre, Seyla Benhabib, and Shlomo Avineri. These scholars called attention to the mistaken assumptions about the nature of the self on which liberal philosophy, especially as espoused by John Rawls, rested. Liberalism, these communitarian critics pointed out, views the person as divorced from all his moral commitments and communal attachments. These communitarians challenged this view, depicting the self as fundamentally "situated" or "contextualized" in a given culture, within a particular history, and with a particular set of values. These academic communitarians, and the sociologists who preceded them, like Émile Durkheim and Ferdinand Tönnies, stressed that individuals in viable communities not only flourish as human beings but also are more reasonable and productive than isolated individuals. Only if social pressures to conform rise to excessively high levels do they undermine the development and expression of the self.

Academic communitarians also argued that the nature of the political community was misunderstood by liberalism. Where liberal philosophers described a neutral framework of rules within which a diversity of commitments to moral values can coexist, communitarians showed that such a "thin" conception of political community was both empirically misleading and normatively dangerous. Good societies, according to these authors, rested on much more than "neutral" rules and procedures; they relied on shared moral culture.

Some academic communitarians argued even more strongly on behalf of particularistic values, suggesting that, indeed, these were the only kind of values that mattered and that it was a philosophical error to posit any universal moral values. Alasdair MacIntyre wrote that human rights are as real as unicorns. Michael Walzer initially argued that concrete universal values were philosophically illusory and that societies could be measured only according to their own particularistic moral standards. As the debate over abstract universal values gave way to a discussion about cross-cultural justifications of human rights, the problems of such a relativistic position came to be widely (though not universally) acknowledged. In the 1990s, responsive communitarians developed a position that accommodated both particularistic and universal values (see below).

It should be noted that, despite being widely referred to as communitarians, Taylor, Sandel, and Walzer systematically avoided the term. Arguably, this was the case because the term used to be, and to some extent still is, associated with authoritarian communitarianism.

Authoritarian Communitarianism

Often referred to as East Asian communitarians, authoritarian communitarians argue that to maintain social order and harmony, individual rights and political liberties must be curtailed. They hold that individuals find their role and meaning in their service to the common good, are organic parts of a great whole, just as human cells are in a human body.

Some East Asian communitarians believe in the strong arm of the state (e.g., former Singaporean Prime Minister Lee Kuan Yew and Malaysian head of state Mahathir Bin Mohammad) and some in strong social bonds and the voice of the family and community (especially the kind of society Japan had, at least until 1990). Authoritarian communitarians argue that the Western value of *liberty* actually amounts to social, political, and moral anarchy and that legal and political rights are a distinctively Western idea that the West uses to impose its own vision on other cultures, which have their own preferred values. Over the years, East Asian communitarians have moderated many of these claims, have made more room for individual rights, and have been increasingly eclipsed, at least in the West, by responsive communitarians.

Responsive (Political) Communitarianism

Early in the 1990s, a new group was founded by Amitai Etzioni, working with William A. Galston, which took the communitarian philosophy from a small academic group, introduced it into public life, and recast its academic content. The group, variously referred to as "responsive," "neo," or "political" communitarians, stressed the importance of society and its institutions above and beyond that of the state and the market, the focus of other public philosophies. It emphasized the key role played by socialization, moral culture, and informal social controls rather than state coercion or market pressures. Responsive communitarianism served as a major correction to authoritarian communitarianism by stressing that strong rights presume strong responsibilities and that one should not be neglected in the name of the other.

The group started by forming a platform, whose drafters included Mary Ann Glendon (law); Thomas Spragens Jr.; James Fishkin and Benjamin Barber (political science); Hans Joas, Phillip Selznick, and Robert Bellah (sociology); and Alan Ehrenhalt (author). The platform was initially endorsed by more than 150 public leaders from across the political spectrum. The voice the group raised was soon found in numerous op-eds and public lectures and on TV and radio programs in a considerable number of countries. The group also issued several position papers on subjects such as organ donation, character education, and HIV testing. In the 1990s, several of the members of this group worked with the New Democrats and advocates of the Third Way in Europe.

The ideas of the group were further developed, both on the public side and on the academic side, in books and in an intellectual quarterly, *The Responsive Community*. These works stressed that social institutions and public policies should reflect shared values and the common good in addition to aggregation of individual preferences, which themselves are culturally penetrated. Beyond universal principles, communitarianism emphasizes particularism, the special moral obligations people have to their families, kin, communities, national societies, and the nascent global community.

Responsive communitarians showed that society is best understood not as composed of millions of individuals, but as pluralism within unity. That is, subcultures and loyalties to various ethnic and regional communities do not undermine the integrity of society as long as a core of shared values and institutions—such as the Constitution and its Bill of Rights, the democratic way of life, and mutual tolerance—are respected. These observations are of special import today for societies that are in the process of coming to terms with mass immigration (e.g., many European societies and Japan) and in which minorities are questioning their place in the national whole (Quebecois, Scots, Basques, Sunnis in Iraq, etc.).

Although this model of the good society is applicable to all societies, at different moments in history, a given society is likely to miss the desired balance between rights and the common good or between particularistic loyalties and society-wide bonds but in a direction different from that of others. Hence, different societies may need to move in different directions in order to approximate the same balance. Thus, contemporary East Asian societies require moving toward much greater tolerance for individualism and pluralism, while in American society, as Robert Bellah and his colleagues along with many others have shown, excessive individualism ought to be reined in.

Applications of Responsive Communitarianism for Political Science

The following are a few illustrative, salient issues to which responsive communitarianism has contributed.

International Relations Theory

Although communitarianism has dealt mainly with national societies, it has also been applied on the international level. Given the rapid increase of pressing transnational issues, such as environmental degradation, illegal immigration, financial crises, transnational mafias, and terrorists with global reach, the need for significant changes in transnational governance (sometimes referred to as global governance) seems evident. As a rule, transnational problems are still tackled by the Old System—by national governments and the international organizations managed by national representatives and funded by national allotments. But in most cases, this Old System has proved to be inadequate to cope with these problems.

Communitarians have addressed the question of how new transnational institutions may be constructed. They argued that just as in early ages, loyalties and commitments once limited to local communities were transferred in part to national communities, in the future, some such transfer will have to take place to global institutions. Supranationality, the concept that forms the backbone of the communitarian theory of international relations outlined here, characterizes a political body that has acquired some of the attributes usually associated with a nation, such as political loyalty and decision-making power, based not on an aggregate of national decisions or those made by representatives of the member states but rather on those made by the supranational bodies themselves. Their capacity to make decisions on their own terms allows supranational bodies to move with much greater agility and speed than international organizations. At this stage, few such institutions exist and those that do face strict limitations because nations still form the major communitarian bodies. A study of the International Criminal Court, the Internet Corporation for Assigned Names and Numbers, and the World Trade Organization illustrates this point. The European Union, the most advanced attempt to form supranational institutions, is not global but regional, and it too, so far, has had only limited success in transferring significant communal loyalties to a transnational level.

Homeland Security

Liberals and especially libertarians tend to seek to protect individual rights against the claims of homeland security. For instance, they call for treating terrorists like other criminals, according them the full rights of citizens. East Asian communitarians tend to hold that security is a common good that trumps individual rights and, hence, support policies that give up civil liberties for the sake of security. Responsive communitarianism holds that there must be a carefully crafted balance between rights and security and that the point of balance changes as technical and international conditions change. They point out that courts in free societies regularly use the terminology of balance between the public interest and individual rights and allow the latter to be curtailed when they undermine a "compelling public interest"—for instance, allowing the violation of privacy of sex offenders in order to protect children from sex abuse and authorizing wire taps for suspected killers. Responsive communitarians have extended the same concept to homeland security.

On reviewing new security measures introduced in the United States after the terrorist attacks of September 11, 2001; in the United Kingdom after the IRA's terrorism; and in Spain after the Madrid train bombings, responsive communitarians found a somewhat more complex picture than the claim that fear mongering had led to wholesale violations of rights ("privacy is dead") or that rights advocates block vital security measures (indeed that even raising such questions is "aiding and abetting the enemy").

Some measures are fully justified, indeed overdue. These often entail a mere adaptation of the law for technical developments. For example, the Foreign Intelligence Surveillance Act (FISA) of 1978 provided guidelines under which a federal agent could obtain authorization to conduct surveillance for "foreign intelligence purposes." Prior to 9/11, wiretap warrants were limited to a given phone. Due to the increasing use of multiple cell phones and e-mail accounts over the last decades, federal officials engaged in court authorized surveillance under FISA could not follow suspects as they changed the instruments they were using unless they got a new court order for each communication device. The USA PATRIOT Act, enacted in October 2001, amended the FISA law to allow what is called "roving surveillance authority," making it legal for agents to follow one suspect, once a warrant is granted, whatever instrument he or she uses. It seems that unless one holds that terrorists are entitled to benefit from new technologies but law enforcement is not entitled to catch up, this is a reasonable measure. (Note that the American constitution bars only "unreasonable" searches and was never absolute in this matter. Although the PATRIOT Act has become a symbol for great excesses in hasty pursuit of security, only a small fraction, about 15 of its more than 150 measures, have been seriously contested. Most are considered reasonable.)

Whatever the new security risks, there are measures that no nation should resort to. These taboo tactics should remain so even if one's adversaries employ them. Banning these measures sets free societies apart from those whose regimes and actions are illegitimate and provides them with the moral high ground. Key examples of such measures are torture, the indefinite suspension of habeas corpus, and mass detention on the basis of race, ethnicity, or national origin.

Further, responsive communitarians hold that many measures are neither inherently justified because enhanced security requires them nor inappropriate because they wantonly violate rights. Instead, their status is conditioned on their being subject to proper oversight. A measure that may seem tilted toward excessive attention to security may be tolerated if closely supervised by oversight organs, while a measure that is judged as tilting toward excessive attention to individual rights may be tolerated if sufficient exceptions are provided that are backed up by second balance organs. That is, new measures can either be excessively privileged (undermining either security or the regime of rights) or excessively discriminated against (leading to inaction on behalf of either element of a sound balance).

The balance sought here is not between the public interest and rights but between the supervised and the supervisors. Deficient accountability opens the door to government abuses of power, and excessively tight controls make agents reluctant or incapable to act.

In effect, the history of homeland protection can be told in terms of lack of accountability being overcorrected, which in turn is overreduced, in the quest for a sound balance. In the United States, the FBI went way overboard in its surveillance of alleged Communists and civil rights activists. In the 1970s, the FBI was reined in by the Church Committee and additional rules the FBI itself issued—rules that in turn became one reason why 9/11 was not prevented. In reaction, the Bush administration introduced a large number of extreme security measures—which were curtailed by the Obama administration as of 2009. Whether the proper balance has been reached between rights and security and the proper level of accountability is far from clear.

Communal Moral Dialogues as Agents of Change

Thus far, this entry has focused on responsive communitarianism's answer to the question of what makes a good society—a balance between social order and individual rights. The remainder of this entry deals with the next question of how understandings of the good are shaped and spread in societies.

One widely shared liberal answer to the question just articulated—what are the main ways and processes by which people's sense of what is legitimate is shaped and reshaped?—is provided by the students of deliberative democracy. These authors point to "reasoned deliberations" as the key way in which citizens of a democracy come to change their judgments and emphasize "cool" and rational processes while minimizing the role played by emotions and other such "hot" factors.

Responsive communitarians argue that such cool, rational deliberations are almost impossible to achieve or even to approximate under most circumstances. The examination of actual processes of decision making, especially when they concern normative matters, shows that they are much more impassionate and proceed by means different from those depicted by the champions of deliberative democracy. More important, deliberative democracy is the "wrong" model for determining the normative bases of political acts—for determining the good. Instead, individual preferences and judgments are shaped largely through interactive communications about values, through "moral dialogues" that combine passion with normative arguments and rely on nonrational processes of persuasion, education, and leadership. According to responsive communitarians, although information and reason have an appropriate role in dialogues about which policies are legitimate, information and reason play a much smaller role than is often asserted. This is so both because they are much weaker tools than believed and because a much greater role is played by another factor: the appeal to values.

Moral dialogues often have no clear opening point or closing event. They are prolonged, heated, and seemingly meandering. However, they often lead to new or reformulated shared normative understandings (cases in point follow). Even very large and complex societies engage in moral dialogues that lead to changes in social definitions of the good. These dialogues take place by linking millions of local conversations (e.g., between couples, in neighborhood bars, in coffeehouses or teahouses, and around watercoolers at work) into society-wide networks and shared public focal points. They take place during regional and national meetings of many thousands of voluntary associations in which local representatives dialogue—in state, regional, and national party caucuses; in state assemblies and in Congress; and increasingly via electronic links (e.g., groups that meet on the Internet). Focal points of these dialogues are national call-in shows, debates on network television, and nationally circulated newspapers and magazines.

Society-wide moral dialogues are often fostered, accelerated, and affected by public events such as hearings (e.g., the Clarence Thomas/Anita Hill case focused discussion on sexual harassment), trials (e.g., the 1925 Scopes trial challenged the normative standing of evolution), and demonstrations (e.g., those that highlighted the normative case against the war in Vietnam).

A brief illustration follows. Until 1970, the protection of the environment was not considered a shared core value in most societies. A book, Rachel Carson's *Silent Spring*, which was very widely read and discussed, triggered a society-wide moral dialogue. A massive oil spill and the ensuing protests and the Three-Mile Island incident further established the subject on the national normative agenda. An estimated 200,000 people gathered on the Capitol Mall in 1970 in Washington, D.C., to demonstrate concern for the environment on Earth Day. These are a few of many such incidents. All these served as dialogue starters and were followed by billions of hours of dialogue, culminating in a shared sense that both persons and communities had a moral commitment to Mother Earth.

The new commitment was strong enough to legitimate various public acts. It moved a conservative president, Richard Nixon, to establish the Environmental Protection Agency and pushed Congress to introduce many policies favorable to the environment, such as recycling and Corporate Average Fuel Economy (CAFE) standards. Similar dialogues led to profound changes in what is considered just treatment of minorities, especially African Americans (dialogues triggered and nurtured by the civil rights movement), relationships between genders (by the women's movement), and many others. Some current dialogues in American society, concerning same-sex marriage and the death penalty, have already affected the normative culture and public policies, though they are less advanced than those listed above.

The opening up of many societies (especially former communist ones), the spread of education, the widening of people's attention horizon from the local toward the national and even global public affairs, the rise of worldwide TV networks (the "CNN effect"), increased travel and immigration, and the World Wide Web—all arising over the past few decades—have led to the development of transnational moral dialogues, in general, and dialogues about what is considered legitimate, in particular. These are global not in the sense that all citizens participate, let alone agree, but in that these dialogues reach across most borders. Thus, concern for the environment now is very widely shared; following transitional moral dialogues, a widely shared understanding evolved that the 2003 invasion of Iraq by the Bush administration was not legitimate; and shared moral understandings have been inching forward on matters as different as the bans on sex tourism, land mines, and whale hunting.

Amitai Etzioni
The Institute for Communitarian Policy Studies
Washington, D.C., United States

See also Civil Society; Republicanism; Social Capital

Further Readings

Avineri, S., & de-Shalit, A. (Eds.). (1992). *Communitarianism and individualism.* Oxford, UK: Oxford University Press

Bell, D. (1993). *Communitarianism and its critics.* Oxford, UK: Clarendon Press.

Bell, D. (2000). *East meets West: Human rights and democracy in East Asia*. Princeton, NJ: Princeton University Press.

Bellah, R., Madsen, R., Sullivan, W. M., Swidler, A., & Tipton, S. (1985). *Habits of the heart.* Berkeley: University of California Press.

Benhabib, S. (1992). *Situating the self: Gender, community, and postmodernism in contemporary ethics.* New York: Routledge.

Ehrenhalt, A. (1995). *The lost city: Discovering the forgotten virtues of community in the Chicago of the 1950s.* New York: Basic Books.

Etzioni, A. (1996). *The new golden rule*. New York: Basic Books.

Etzioni, A. (2004). *From empire to community.* New York: Palgrave Macmillan.

Frazer, E. (1999). *The problems of communitarian politics*. Oxford, UK: Oxford University Press.

Glendon, M.-A. (1991). *Rights talk: The impoverishment of political discourse.* New York: Free Press.

The Institute for Communitarian Policy Studies. (1991). *The responsive communitarian platform*. Retrieved November 2, 2010, from http://www.gwu.edu/~ccps/rcplatform.html

MacIntyre, A. (1984). *After virtue.* Notre Dame, IN: University of Notre Dame Press.

Mulhall, S., & Swift, A. (1992). *Liberals and communitarians.* Oxford, UK: Blackwell.

Obama, B. (2006). *The audacity of hope.* New York: Crown/Three Rivers Press.

Sandel, M. (1981). *Liberalism and the limits of justice.* Cambridge, UK: Cambridge University Press.

Selznick, P. (2002). *The communitarian persuasion.* Washington, DC: Woodrow Wilson Center Press.

Taylor, C. (1989). *Sources of the self: The making of the modern identity.* Cambridge, UK: Cambridge University Press.

Walzer, M. (1983). *Spheres of justice.* Oxford, UK: Blackwell.

COMPARATIVE METHODS

At the most general level, the term *comparative methods* may be used to refer to any research technique that focuses on patterns of similarities and differences across cases. More commonly, however, the term is used to refer to a family of techniques employed by small-*N* researchers. It is this family of techniques that is the main focus of this entry. Though comparative methods do have particular advantages for small-*N* researchers, there is no inherent connection between comparative methods and small *N*s. The development of formal methods of comparative research—specifically, qualitative comparative analysis (QCA) and related techniques—has facilitated the extension of comparative logic, which is configurational in nature, to large-*N* studies.

This entry has six major sections. The first emphasizes the case-oriented nature of comparative research. This feature clearly differentiates it from conventional variable-oriented research and underscores its broad links to qualitative inquiry in general. The second section considers the goals of comparative research, linking comparative research to both inductive theory building and deductive theory testing. The logic of comparative analysis is the central focus of the third section. The importance of the analysis of set-theoretic relations to comparative research is emphasized, and the analysis of sufficiency and necessity is linked to subset and superset relations. These connections are important because much of social science theory is formulated in terms of set relations, and it is important to use analytic methods that are true to theoretical formulations. The language and logic of sets underlie both informal and formal methods of comparative analysis, which is the topic of the fourth section. Traditionally, comparative researchers have made use of informal methods, explicitly or implicitly relying on John Stuart Mill's methods of agreement and difference. Over the past 2 decades, Charles Ragin has formalized the comparative method as QCA, which is discussed in detail in this section. In the fifth section, discontinuities between comparative research and conventional variable-oriented research are addressed, and in the closing section, current development in formal methods of comparative analysis is described.

The Case-Oriented Nature of Comparative Research

Comparative researchers tend to conceptualize causality in terms of necessity and sufficiency. A common concern is the combination of conditions that are sufficient for an outcome. Often, comparative researchers find that different combinations of causal conditions generate the same outcome. This approach to causal conditions is distinct from that of variable-oriented research programs, which seek to identify the net effect of each independent variable on the dependent variable, isolating its effect from that of competing variables.

While the sufficiency of causal combinations is a major preoccupation of comparativists, necessary conditions are also important, especially to theory. For example, Barrington Moore's famous maxim, "No bourgeoisie, no democracy," exemplifies the necessary condition: The rise of a capitalist class is necessary for—but by no means guarantees—the development of democracy. The bourgeoisie would also need to resist domination by the aristocracy and peasantry, promote the development of commercial agriculture, and participate in a revolutionary break with the past. This combination of conditions, Moore found, was sufficient to produce a democratic state in the cases he studied.

The analysis of necessity and sufficiency requires that cases be considered on their own terms, as interpretable wholes. To identify and make sense of the bourgeoisie's revolutionary role, Moore conducted a series of historical case studies. For each case, Moore examined how various aspects of the case (e.g., the relationships among the bourgeoisie, aristocracy, and peasantry and the development of commercial agriculture) interacted with and influenced each other. This is the defining characteristic of case-oriented research: A case constitutes a coherent whole, and its various aspects are understood relationally. Thus, case-oriented research is inherently *configurational.* The comparative aspect of comparative research, therefore, is the comparison of configurations of interconnected aspects—the cross-case analysis of configurations of similarities and differences.

Within the social sciences, comparative research is most commonly associated with small-*N* macrocomparative research. But in fact, its logic may be applied at any level of analysis. The same techniques that a macrolevel comparative researcher might use to examine the relationship between economic development and the emergence of democracy could be used by a microlevel comparative researcher exploring the relationship between religious beliefs, socioeconomic status, and divorce. What matters is not the number of cases under investigation but that the researcher make sense of these cases by treating them as instances of interconnected aspects, respecting their coherence and integrity as cases.

There is, of course, always a trade-off between the breadth and the depth of an analysis, and it is the researcher's responsibility to determine whether a given research question is best answered by examining fewer cases in great detail or more cases in less depth. Generally, this choice is dictated by the goals of the research.

The Goals of Comparative Research

Comparative methods can be used for either theory development or evaluating hypotheses. When developing theory, the researcher conducts a series of case studies, strategically selecting cases so as to produce an accumulation of knowledge. For example, a specific type of evidence that is not available in one case may prompt the investigator to select for his or her next study a case that offers this evidence. Questions raised by previous case studies direct the investigator's subsequent case selection. In this manner, comparative researchers can develop their theoretical arguments inductively, building their theories from the bottom up. This grounded approach of theory development can be contrasted with a deductive, hypothesis-testing approach that begins with multiple cases. Here, the researcher starts with a preliminary specification of the relevant cases. For example, a researcher interested in testing Theda Skocpol's theory of social revolution would begin by identifying cases that fall within her specified scope conditions, as explained by Gary Goertz in his book *Social Science Concepts: A User's Guide*. The researcher would then examine the characteristics of these cases to see if they are consistent with Skocpol's theory—namely, that countries experiencing social revolutions exhibit the relevant causal conditions (e.g., state breakdown combined with peasant revolts and some form of international pressure).

The inductive approach to comparative research seeks to develop an individual explanation for each case, in relative isolation from other cases. Of course, it is impossible for researchers to wear blinders or forget what they have already learned, and knowledge of one case will invariably influence one's understanding of subsequent cases. Nor would such a "clean-slate" approach be desirable—the knowledge gained during the research process assists the researcher in determining where to focus his or her attention next. Still, the goal of the inductive approach is to produce a series of individual case studies, allowing for maximum diversity. It is only on completing the individual case studies that the researcher turns to the task of making an encompassing comparison. The researcher examines each case against the others, sometimes by applying formal techniques, such as QCA. The goal of this cross-fertilization of case studies is to identify what the cases do and do not have in common. The researcher uses these comparisons to develop an analytic frame that clarifies the nature of the cases under investigation and their relationship(s) to one another. The construction of this analytic frame produces a structured theoretical account of the cases that may be used to identify cross-case patterns of similarity and difference (e.g., different paths to the same outcome).

Where the inductive approach concludes with the development of an analytic frame, the deductive approach begins with it. Based on existing theoretical and substantive knowledge, the researcher develops an analytic frame to guide his or her investigation of the hypothesis that he or she intends to test. Here, think of an analytic frame as akin to a survey instrument that "interrogates" each respondent (case) in a uniform manner. The analytic frame directs the researcher's investigation, identifying plausible cases for inclusion in the analysis and telling him or her which characteristics of the cases to examine, what questions to ask, and how to record the information. Fundamental to the deductive approach are the goals of unbiasedness and efficiency, both of which a clearly defined analytic frame helps achieve. Having completed data collection, the researcher then seeks to assess the hypothesis. As in inductive comparative research projects, a formal technique such as QCA may be used to simplify the analysis. In fact, it is somewhat more common for QCA to be used in deductive comparative research than in inductive research because *N*s tend to be larger.

The inductive and deductive approaches are combined when a researcher seeks to study a theoretically specified range of cases. As with the deductive approach, the researcher relies on a preexisting analytic frame to guide case selection. Here, the analytic frame serves as a device for classifying cases according to aspects of theoretical or substantive interest. For example, the researcher might construct a contingency table cross-tabulating the level of development of countries with their level of democracy. The researcher then seeks to fill each cell of this table with at least one case. He or she can then subject these cases to in-depth, inductive analysis, confident in the knowledge that important case aspects are well represented in the study.

Inductive comparative research proceeds in a bottom-up fashion, accumulating case knowledge so as to generate new theory. By contrast, deductive comparative research follows a top-down formula. Beginning with theory, the deductive comparative researcher seeks out cases to subject to some form of hypothesis testing. Of course, neither of these strategies is practiced in pure form, and there are many examples where the two are explicitly blended. Regardless of the direction in which the research proceeds, however, the same basic analytic logic is employed.

The Logic of Case-Oriented Comparison

As noted previously, comparative research is fundamentally case oriented. Cases are conceived as configurations, and comparison of these configurations is achieved by categorizing cases into sets. Thus, the logic of comparative research is fundamentally set theoretic. Set-theoretic social research, based on Boolean algebra, involves the analysis of set relations, especially the identification of subset relationships. Most social theoretical statements are framed in terms of subset relationships. For example, when social researchers observe that "religious fundamentalists are politically conservative," they are specifying a subset relationship: Religious fundamentalists constitute a subset of political conservatives. The observation that economically developed countries are democratic follows the same pattern; the conjecture is that the former constitute a subset of the latter.

Subset relationships such as these—ones in which the cause (e.g., religious fundamentalism or economic development) is a subset of the outcome (e.g., political conservatism or democracy)—are consistent with a relationship of sufficiency. That is to say, membership in the set of religious fundamentalists (economically developed countries) is sufficient for membership in the set of political conservatives (democratic countries). Conditions of necessity, by contrast, are identified by establishing that the outcome is a subset of the cause. Referring again to Moore's maxim, which may be understood as the specification of a subset relationship, democratic countries (outcome) constitute a subset of countries with a bourgeois class (causal condition).

Of course, subset relationships need not be perfect. For example, there are certainly some religious fundamentalists who are politically liberal. Some relationships, therefore, might be more accurately characterized as "rough" subsets. Ragin has proposed a measure of set-theoretic consistency that assesses the degree to which a given set of evidence conforms to the subset relationship. Clearly defined subset relationships produce high consistency scores; as consistency drops, the claim that a rough subset relationship exists is less tenable.

While having a measure of set-theoretic consistency is useful for assessing subset relationships empirically, researchers ultimately must rely on their theoretical and substantive knowledge (which includes their knowledge of the cases under investigation) when determining whether the claim that a subset relationship exists—even a rough one—is justified. Social life is complex, and perfect subset relationships are rare, especially when the number of cases is large. Individuals, for example, are notoriously inconsistent. Although it is highly unlikely that an individual with little education will achieve a high-prestige occupation, it has happened. Whether such occurrences are considered common enough to challenge the claim that a rough subset relationship exists between education and occupational prestige is a determination that the researcher must make. Is a high level of education "almost always necessary" for a high-prestige occupation? Such questions cannot be answered through the rote execution of a consistency calculation but instead require that researchers apply their theoretical knowledge and interpretive skills.

The analysis of set-theoretic relations has other uses, in addition to the assessment of conditions of necessity and sufficiency. Set-theoretic analysis also facilitates theoretical development by providing an empirical basis for constructing and evaluating ideal types. Max Weber conceived of the ideal type as a "one-sided exaggeration" that serves to distinguish a class of social phenomena. The usefulness of the ideal type, Weber argued, is that it is not designed to perfectly represent all aspects of the phenomenon under investigation but, rather, its essential aspects. Once these essential features have been established, it is possible to assess, using set-theoretic methods, the degree to which empirical cases conform to the ideal-typical formulation.

The categorization of cases according to their degree of memberships in sets (using fuzzy sets) is crucial to assessing their degree of conformity to an ideal type. Gøsta Esping-Andersen's distinction among types of welfare regimes is a case in point. There are, of course, no pure cases of liberalism, social democracy, or corporatism. In practice, states of each type borrow from the others. How, then, does one decide on the proper categorization of welfare states? The first step is to conceive of each welfare regime as an ideal type and to specify the essential elements that constitute it as a type. These elements can then be formulated as sets, and the degree of membership of each empirical case in each set can be assessed (via fuzzy-set membership scores). The membership of each empirical case in each ideal-typical regime type is given by the minimum of its memberships in the constituent sets for each regime type. (Using the minimum in this way is known as fuzzy-set intersection.) The use of the minimum is based on the simple idea that degree of membership of an empirical case in an ideal type is only as strong as its weakest link (lowest membership score), a principle that follows from the theoretical notion of the ideal type as a specific combination or coalescence of essential elements. The analysis of necessary and sufficient conditions and the construction of ideal types are not necessarily independent of one another. Frequently, the results of one analysis will prompt the researcher to reconsider the results of the other. Whether they recognize it or not, comparative researchers commonly engage in both forms of analysis.

Formal Versus Informal Methods of Comparative Research

Classical studies in comparative political economy such as Barrington Moore's *Social Origins of Dictatorship and Democracy* and Theda Skocpol's *States and Social Revolutions* rely on informal methods of comparative analysis, and to date, most comparative researchers continue in this tradition. It was with the publication of *The Comparative Method* in 1987 that formal methods of comparative analysis first gained prominence. Informal and formal comparative methods share an underlying logic and, in the hands of a capable researcher, will produce the same general conclusions. The difference is that formal methods explicitly make use of Boolean algebra to structure the analytic process.

The logic of comparative analysis and inductive inquiry in general can be traced to the work of John Stuart Mill and George Boole. In *A System of Logic*, Mill proposed a series of methods of inductive reasoning that would lay the logical foundation of informal methods of comparative research. Contemporaneously, Boole was developing his calculus of truth values. Boolean algebra, as it came to be known, would develop into set algebra and, ultimately, be integrated into formal methods of comparative research, especially QCA.

Mill's Method of Agreement and Indirect Method of Difference

The simplest of all comparative logics are Mill's method of agreement and his indirect method of difference, both of which assume that any given outcome has one and only one cause. Although Mill described these methods in terms of individual causes, it is straightforward to extend these methods to accommodate combinations of causal conditions when it makes sense to conceptualize them as decisive causal conjunctions. A researcher using the method of agreement examines a set of cases, each of which exhibits the same outcome. If the researcher finds that all the cases share one or more causal conditions, he or she can infer that this causal condition (or combination of causal conditions) is linked to the presence of the outcome. The indirect method of difference operates under related premises, except that both positive and negative instances of the outcome are examined. This method applies agreement and difference *seriatim*; for this reason, it is often referred to as the joint method of agreement and difference. The researcher first applies the method of agreement to cases exhibiting the outcome. On identifying a causal condition (or causal combination) linked to the outcome, the researcher examines the negative cases to see if the absence of the outcome is associated with the absence of the cause or causal combination. This double application provides researchers with stronger evidence on which to base their conclusions.

Comparative researchers continue to employ these methods today, albeit under different monikers. The method of agreement is the basis of the most different systems design, which is better conceived as a most different/same outcome design (MDSO). This alternate labeling makes it clear that in this design, diverse cases share the same outcome.

The researcher attempts to identify the decisive, causally relevant conditions shared by these diverse cases. The other popular design implements the indirect method of difference. It is commonly known as most similar systems design, but it is better understood as a most similar/different outcome design (MSDO), making it clear that there are both positive and negative cases. Ideally, negative cases should be matched as closely as possible with positive cases.

Mill's methods also provide the underlying logic of the "most likely," "least likely," and "crucial" case study research methods. The most likely case study employs cases that theory predicts should strongly exhibit the outcome, while the least likely case study employs cases that would do so weakly, if at all. Least likely cases that conform to predictions serve to marshal support in favor of a given theory. In contrast, most likely cases that do not behave as predicted cast doubt on a theory's validity. The strongest claim is made by the crucial case study, which asserts that a case must exhibit the outcome if the theory is not to be falsified.

The limitation of Mill's methods and derivative techniques is that they are incapable of addressing various forms of causal heterogeneity, especially equifinality and multiple conjunctural causation. Mill himself noted that his techniques fail when an outcome has more than one cause. Researchers are able to overcome this deficiency by immersing themselves in the cases. This immersion, however, necessarily restricts the breadth of analysis. It is perhaps for this reason that comparative research has become so closely associated with small-*N* analysis. QCA challenges this affinity by formalizing many aspects of the comparative research process and enables comparative researchers to address larger *N*s.

Qualitative Comparative Analysis

QCA was originally developed for use with binary data or "crisp sets." The logic of QCA, however, extends to fuzzy sets in which set membership scores can be coded in the interval from 0.0 to 1.0. The present discussion encompasses both types of data. Note that crisp-set analysis is easier for novices to grasp, while fuzzy-set analysis provides more nuanced results. There are four basic steps in most applications of QCA. In practice, of course, these are intimately connected, and researchers should view the research process as a dialogue among four analytic moments. In other words, comparative research is usually retroductive rather than predominantly inductive or deductive.

Step 1: Calibration

As has been emphasized, comparative research is fundamentally set theoretic. The classification of cases into sets, therefore, is of primary importance. This classification process is referred to as calibration. Calibration is a routine practice in the natural sciences. Simply put, calibration refers to the practice of adjusting measuring instruments so that they conform to dependable, known standards. The act of adjusting one's watch is an act of calibration as is checking the accuracy of a thermometer. While calibration is routine in the natural sciences, most social scientists make use of uncalibrated measures that merely show the positions of cases relative to one another. Uncalibrated measures, however, are clearly inferior to calibrated measures. With an uncalibrated thermometer, for example, it is possible to know that one object has a higher temperature than another, but one still does not know whether either object is hot or cold. This is the key: Calibrated measures provide directly interpretable values. Calibration permits researchers to classify a country as democratic or autocratic, not merely more versus less democratic than the next (or the average).

The reason that crisp-set analysis is more accessible is because the calibration process is more straightforward. When calibrating crisp sets, the researcher need only determine whether each case is part of the target set or not. So, for example, a poverty researcher might seek to classify respondents as "poor" or "not poor." When calibrating fuzzy sets, by contrast, the researcher must determine the degree to which each case belongs to the target set. The relevant distinctions here might be between respondents who are poor (membership = 1.0), somewhat poor (membership = 0.67), not that poor (membership = 0.33), and not poor (membership = 0). As with a calibrated thermometer, a calibrated fuzzy set integrates both qualitative and quantitative assessments in a single instrument.

Regardless of whether the researcher is calibrating crisp or fuzzy sets, the process demands that he or she carefully consider the nature of the target set—specifically, what the target set represents and how it is labeled. In other words, conceptualization and calibration are closely intertwined. With crisp sets, a case is either in or out of the target set, and sometimes, cases must be force-fitted into categories. Alternatively, the researcher may generate new crisp categories to capture wayward cases. Fuzzy-set calibration, by contrast, permits the use of finely grained distinctions. When calibrating fuzzy sets, it is important for the researcher to specify thresholds for set membership and nonmembership in the set, as well as what is known as the crossover point (fuzzy membership = 0.5), which distinguishes cases that are "more in" the target set from those that are "more out." The calibration process concludes with the production of a data set. As with a conventional data set, cases occupy the rows. The columns, however, do not represent conventional variables but instead represent sets, and the values in the cells indicate the degree to which each case belongs to each set.

Step 2: Necessity Analysis

Having generated the calibrated data set, the researcher next tests for the presence of necessary conditions. As described earlier, the necessity test takes the form of a set-theoretic analysis. A necessary causal condition is one that is a superset of the outcome. Visually, a set relation consistent with necessity is represented by a scatterplot of two fuzzy sets (with outcome membership scores on the *Y* axis and membership in the causal condition on the *X* axis) in which all cases reside in the lower-right triangular region. With a calibrated data set, this situation can easily be tested, and a relationship of necessity can be established if, for each case, the value of the causal condition is greater than or equal to the value of the outcome (i.e., $X_i \geq Y_i$), within the bounds of the consistency threshold. This test can be extended to combinations of causal conditions, and software packages such as fuzzy-set QCA (fsQCA) contain functions to automate the procedure.

Step 3: Truth Table Construction

Truth tables are used to assess the sufficiency of causal conditions, especially in combinations. Necessary conditions are excluded from the truth table analysis, which is why the researcher tests for them first. Sufficiency is indicated when a causal condition (or, more commonly, a combination of causal conditions) constitutes a consistent subset of the outcome and is mathematically represented by the equation, $X_i \leq Y_i$. Again, the consistency test is conducted within the bounds of the consistency threshold established by the researcher.

Superficially, a truth table appears similar to a conventional data set in that it presents a cases-by-variables format. But the rows of a truth table are not observations as they are in a conventional data set. Rather, each row represents a logically possible combination of causal conditions (which, in turn, can be seen as constituting an ideal-typical case or configuration). Consisting of one row per logically possible combination of conditions, a truth table has 2^k rows, where k equals the number of causal conditions. As the number of causal conditions increases, the size of the truth table grows exponentially, and analysis becomes increasingly complex. Software packages help manage this complexity; nevertheless, most practitioners restrict their analysis to between 4 and 10 causal conditions.

To construct the truth table, the researcher refers to the calibrated data set and uses the consistency measure to determine which combinations of causal conditions are linked to the presence of the outcome and which are linked to its absence. With a large number of cases and/or causal conditions, this process can become error prone. Software packages facilitate the conversion of a calibrated data set into a truth table and should be used for all but the simplest analyses. Researchers should keep in mind that the process of converting the calibrated data set into a truth table will frequently force them to reconsider the measurement and calibration of both their causal conditions and outcome. In many ways, the process of constructing the truth table is the heart of the comparative research process, and it must not be approached mechanically. It is during this phase of research that the researcher refines and, ultimately, finalizes his or her analytic frame. Indeed, for a descriptive project, the researcher may choose to conclude simply with a presentation of the truth table. Functionally, the truth table is a compact way of presenting a multidimensional typology.

Step 4: Truth Table Reduction

The analysis of a truth table involves a process known as truth table reduction or Boolean minimization. This process results in a Boolean equation that expresses the various causal combinations that are linked to the presence of the outcome. Researchers interested in the mechanics of this process are directed to Ragin's three books on this subject: *The Comparative Method*, *Fuzzy-Set Social Science*, and *Redesigning Social Inquiry*. The following discussion focuses on the results of the reduction procedure. The truth table reduction produces two main types of solutions, known as the complex and parsimonious solutions, depending on whether or not the researcher chooses to make use of remainders. Remainders are rows of the truth table that lack cases (empirical instances). Remainders are not uncommon. The social world is characterized by limited diversity, and many combinations of causal conditions simply do not exist.

Remainders are best understood as potential counterfactual cases—combinations of conditions that do not exist empirically, at least not in the investigator's data set, and are thus subject to conjecture. By including remainders in the truth table reduction process, the researcher is able to use QCA to conduct counterfactual analysis. The most common use of remainders in QCA is to produce a parsimonious solution in addition to the usual complex solution. The parsimonious solution allows the use of any remainder, as along as its inclusion in the solution results in a logically simpler formula for the outcome. The complex solution bars the use of remainders altogether on the assumption that all the different combinations that are linked to the outcome in question are represented in the investigator's data set.

QCA also provides researchers the opportunity to use their theoretical and substantive knowledge to produce intermediate solutions. This knowledge is input into the software via a simple table and is used to bar the inclusion of untenable counterfactuals from the solution. In general, intermediate solutions are best because they more closely mimic the use of counterfactual cases in traditional case-oriented comparative research. Of necessity, complex solutions constitute subsets of intermediate solutions, and intermediate solutions constitute subsets of parsimonious solutions. In other words, the three solutions are logically nested, according to the degree to which each incorporates counterfactual combinations. Complex solutions incorporate no counterfactual combinations and thus are often needlessly complex; parsimonious solutions are the simplest, but they often incorporate counterfactual combinations that contradict existing theory and knowledge. Intermediate solutions, by contrast, are based on counterfactual analysis that is guided and constrained by theoretical and substantive knowledge.

As has been repeatedly emphasized throughout this discussion, the process of comparative research is iterative. It is likely that the results of the truth table reduction will motivate the researcher to revisit one or more of the previous steps. It is also quite common for researchers to reconsider not simply the calibration process but also whether they are, in fact, capturing the appropriate causal conditions and outcome. Comparative researchers regularly reflect on the nature of their research questions and frequently refine their questions as the research proceeds. QCA does not eliminate the necessity of this process. On the contrary, it makes it explicit. Both overly simplistic solutions and overly complex ones serve as warnings, directing the researcher to reconsider their solutions, revisit their evidence, and sometimes reformulate their analytic frames.

Discontinuities Between Comparative and Variable-Oriented Research

It is best to understand comparative analysis as a distinct methodological approach, one that can be conducted using either informal or formal methods. Far too often, QCA is viewed as a methodology in and of itself when, in fact, it is a formalization and extension of case-oriented comparative research. As has been emphasized throughout this discussion, the comparative approach is fundamentally case oriented, and as such, its tools are designed to answer questions that differ from those of conventional variable-oriented research. Nevertheless, the distinctiveness of comparative research is often overlooked, with the result that comparative researchers and variable-oriented researchers often talk past each other when discussing their research.

Ironically, the popularity of QCA has, to some degree, exacerbated this issue. Just as the rise of desktop statistical software has made it easy to estimate a poorly designed regression model, it is now straightforward for naive users to conduct a QCA without recognizing the different assumptions that they are taking on board. This does not refer to philosophical debates regarding epistemological and ontological differences between case-oriented and variable-oriented strategies. Rather, there are substantial differences between what the research tools themselves produce in terms of social-scientific representations.

The importance of these differences can be seen clearly in the contrasts between how each perspective establishes causality. As this entry has discussed, comparative researchers establish causality by identifying subset relationships: "No bourgeoisie, no democracy." But such a claim does not make sense from the variable-oriented perspective. Subset relationships typically exhibit little to no covariation, which is the foundation for the variable-oriented researcher's causal inference. Similarly, the covariation of variables is insensitive to subset relationships—the type of explicit connection that is central to case-oriented comparative analysis. From each perspective, the other's claim to an empirical basis for causal inference appears flawed. Although both approaches use theory and data to construct representations of social life, they do so in ways that involve sharply contrasting orientations regarding the consideration of cases, the relationship between causal conditions and outcomes, and the criteria of a successful analysis.

The Consideration of Cases

From the perspective of a comparative researcher, the primary problem with variable-oriented methods is that they veil cases. When a variable-oriented researcher, for example, computes a correlation between two variables, it is the variables that take center stage. The cases essentially disappear. In contrast, it is the cases that take center stage in a comparative analysis. It is a simple matter, for example, to map cases onto the rows of a truth table, and some software packages automate this procedure. The two approaches also differ in their methods of case selection. As previously noted, comparative researchers choose cases purposefully, based on either theoretical or substantive criteria. Further, the set of cases included in an investigation may shift as the researcher decides that one or more cases do not "fit" with the others or realizes that an important type of case is missing from the analysis. Variable-oriented researchers, by contrast, typically delineate their samples of cases at the outset of their research. What matters most is that the cases are selected in a manner that ensures that the sample is representative of the larger population, however conceived and defined.

The Connection Between Causal Conditions and Outcomes

Variable-oriented research is usually concerned with explaining variation in the values of a dependent variable—an outcome that varies across cases. The goal of such research is to explain how changes in the values of independent variables affect the value of the dependent variable. Techniques such as linear regression accomplish this by identifying the net contribution of each independent variable. Comparative research, in contrast, is concerned with how causal conditions combine to produce a particular outcome. The solutions of a QCA application, for example, represent different paths or recipes for a given outcome.

In a variable-oriented study, causation is inferred from patterns of covariation. Usually, the researcher measures the relative strength of several causal variables simultaneously. Frequently, the goal is to assess the relative strength of competing theories by determining which independent variable (or set of independent variables) explains the most variation in the outcome variable. In other instances, the goal may be simply to appraise the relative importance of the different independent variables or simply to establish that the causal impact of a theoretically important variable is significant and can be estimated. In virtually all types of variable-oriented analysis, variables compete with one another to explain variation. The moniker of *independent* is important here. Each independent variable is considered on its own terms, capable of affecting the outcome variable regardless of the presence or level of the other causal variables. The comparative researcher, by contrast, does not assume that causal conditions are independent of one another. Independence is the exception rather than the rule, and causation is understood to be conjunctural in nature. The goal of the investigation is to identify combinations of causal conditions that are shared by sets of cases and, in tandem, produce the outcome.

The Criteria of a Successful Analysis

It is useful as well to contrast the different criteria used to judge a successful conclusion of a study. For the variable-oriented researcher, the analysis is successful when a theory can explain why changes in the values of the independent variables affect the value of the dependent variable in a particular manner. For the comparative researcher, by contrast, a successful conclusion is one in which the various paths to the outcome are clearly and convincingly articulated and cases can be associated with the different paths. In short, to justify their conclusions, comparative researchers point to their cases, while variable-oriented researchers point to theory.

From the perspective of variable-oriented researchers, comparative research is replete with analytic sins and errors. Samples are purposively constructed, and their sizes are small. Researchers may add or drop cases at any stage of the analysis. Important causal factors do not vary. In a variable-oriented analysis, any of these situations would undermine attempts at causal inference. Likewise, common practices in variable-oriented research would seriously undermine the validity of the conclusions when viewed from the perspective of comparative research. There are too many cases for an in-depth analysis to be conducted. A fixed sampling frame assumes a level of comparability among cases that may be unwarranted. And the focus on independent effects of variables violates the configurational assumption of comparative analysis. Comparative research is fundamentally case oriented, and consequently, its practices fundamentally clash with those of variable-oriented research. It is therefore not surprising that researchers frequently talk past one another.

Current Developments in Comparative Research Methods

The use of informal methods of comparative research is, of course, well established among social researchers, as is the use of crisp-set QCA. The extension of QCA to encompass fuzzy-set analysis has been welcomed by researchers, although there is a lag between new developments in the approach and their adoption by researchers. For example, the inclusion algorithm (using the principle of set inclusion) described by Ragin in *Fuzzy-Set Social Science* has been superseded by a hybrid algorithm that uses truth tables to synthesize the results of the analysis of fuzzy-set relationships. Although the transition to the truth table algorithm has been widely documented, a number of researchers continue to rely on the procedures discussed in the 2000 text. The publication of Ragin's *Redesigning Social Inquiry* establishes the superiority of the truth table algorithm.

The popularity of QCA has encouraged a number of researchers to propose various extensions to the method. Many of these proposed extensions can be found in working papers at the COMPASSS website, which serves as a hub for researchers interested in the development of formal methods of comparative analysis. Although a number of these proposals aim to incorporate variable-oriented techniques into QCA, such proposals have not gained much traction. The most promising of these extensions are those arising from or inspired by the concerns of traditional small-*N* researchers. Particularly promising is the application of QCA to the analysis of scope conditions, which are used to distinguish between cases that do and do not apply to a given theory. Also promising is the application of QCA to possibility analysis. A relatively recent development, the analysis of possibility seeks to identify those cases that are within a theory's scope but for which the outcome is nevertheless impossible or virtually impossible. The question of whether an individual born into poverty can acquire a high-status occupation, for example, is a question of possibility: What conditions serve to make social mobility possible? The appeal of extending QCA to the study of scope and possibility is easy to see: Such questions are set theoretic in nature. They are theoretically important questions but outside the grasp of conventional quantitative methods.

The popularity of multimethod research suggests that researchers will increasingly seek to use QCA in tandem with other analytic methods. Such projects have proved to be most successful when QCA is combined with both case-oriented and variable-oriented techniques. The combination of QCA and techniques of social network analysis have also proved to be particularly fruitful, and development in this direction is expected to continue. Initiatives to implement the QCA techniques in statistical environments such as R and Stata have begun. By and large, these translation efforts are incomplete, and the canonical package remains

Ragin's fsQCA. Cronqvist's Tools for Small-*N* Analysis (TOSMANA) is also popular. Although lacking fuzzy-set capabilities and procedures for counterfactual analysis and intermediate solutions, TOSMANA includes a number of attractive features such as a threshold-setting tool—to aid the construction of crisp and multivalue sets—and the ability to graphically represent Boolean solutions as Venn diagrams.

The popularity of formal comparative methods is expected to continue to grow. There are two reasons for this. First, QCA permits comparative researchers to include more cases in a given study. Research projects previously too complex to conduct due to the limitations of informal comparative methods are now feasible. Second, and perhaps more important, the process of formalizing the comparative approach served to highlight the fact that despite the popularity of variable-oriented methods, most social theory is actually set theoretic in nature. Many research questions are actually more appropriately answered using comparative methods, because of their set-theoretic nature, than variable-oriented ones. Researchers who previously would have turned to variable-oriented techniques are more likely today to recognize the value of comparative methods.

Charles C. Ragin and Claude Rubinson
University of Arizona
Tucson, Arizona, United States

See also Boolean Algebra; Case Studies; Causality; Conditions, Necessary and Sufficient; Configurational Comparative Methods; Fuzzy-Set Analysis; Qualitative Comparative Analysis

Further Readings

Abbott, A. (2004). *Methods of discovery: Heuristics for the social sciences*. New York: W. W. Norton.

Brady, H. E., & Collier, D. (2004). *Rethinking social inquiry: Diverse tools, shared standards.* Lanham, MD: Rowman & Littlefield.

Breiger, R. L. (2009). On the duality of cases and variables: Correspondence analysis (CA) and qualitative comparative analysis (QCA). In D. Byrne & C. Ragin (Eds.), *Handbook of case-based methods* (pp. 243–259). London: Sage.

Byrne, D., & Ragin, C. (2009). *Handbook of case-based methods.* London: Sage.

COMparative methods for the advancement of systematic cross-case analysis and small-n studies. Retrieved November 10, 2010, from http://www.compasss.org

Eckstein, H. (1992). Case study and theory in political science. In H. Eckstein (Ed.), *Regarding politics: Essays on political theory, stability, and change* (pp. 117–173). Berkeley: University of California Press.

Esping-Andersen, G. (1990). *The three worlds of welfare capitalism.* Princeton, NJ: Princeton University Press.

Fearon, J. D. (1991). Counterfactuals and hypothesis testing in political science. *World Politics, 43,* 169–195.

George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge: MIT Press.

Goertz, G. (2006). *Social science concepts: A user's guide.* Princeton, NJ: Princeton University Press.

Mill, J. S. (1875). *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation* (Vol. 2). London: Longmans, Green, Reader, & Dyer.

Moore, B. (1966). *Social origins of dictatorship and democracy: Lord and peasant in the making of the modern world*. Boston: Beacon Press.

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies.* Berkeley: University of California Press.

Ragin, C. C. (2000). *Fuzzy-set social science.* Chicago: University of Chicago Press.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago: University of Chicago Press.

Ragin, C. C., & Becker, H. S. (1992). *What is a case? Exploring the foundations of social inquiry.* Cambridge, UK: Cambridge University Press.

Ragin, C. C., Berg-Schlosser, D., & De Meur, G. (1996). Political methodology: Qualitative methods. In R. Goodin & H.-D. Klingemann (Eds.), *New handbook of political science* (pp. 749–768). New York: Oxford University Press.

Ragin, C. C., & Rubinson, C. (2009). The distinctiveness of comparative research. In T. Landman & N. Robinson (Eds.), *The SAGE handbook of comparative politics* (pp. 13–34). London: Sage.

Rihoux, B., & Ragin, C. (2008). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. Thousand Oaks, CA: Sage.

Skocpol, T. (1979). *States and social revolutions: A comparative analysis of France, Russia, and China.* Cambridge, UK: Cambridge University Press.

Tilly, C. (2006). *Big structures, large processes, huge comparisons.* New York: Russell Sage Foundation.

Comparative Politics

Comparative politics is a subdiscipline of political science. The goal of political science is to promote the comparison of different political entities, and comparative politics is the study of domestic politics within states. It differs from the other subdiscipline of political science—international relations—which instead focuses on politics between states. Traditionally, it has been assumed that whereas comparative politics studies politics in contexts where there is an ordering principle (the sovereign state), international relations, instead, studies politics in contexts without such a principle (the international system). The first is interested in studying politics in stable domestic contexts, the second in studying politics in unstable, extradomestic contexts. The first has concerned itself with studying order (because it is guaranteed by the sovereignty of the state), the second with studying disorder (an outcome of the anarchy of the relations between states). Some have questioned whether such a distinction between these subdisciplines is still plausible at the beginning of the 21st century.

This entry is divided into four parts, beginning with an analysis of the main methods of comparative politics. The second part discusses the main theories of comparative politics, and the third part identifies some of the issues of comparative politics investigated in the various regions of the world. The entry concludes with a discussion of the future of comparative politics in a globalized world.

Methods in Comparative Politics

Although comparative politics is defined primarily by the phenomena it researches, it is also characterized by the method employed in that research. One cannot engage in comparative analysis without a method for comparing; a method is necessary for testing empirical hypotheses about relations between variables in different cases. Such hypotheses concern the relations between the variables that are held to structure the political phenomenon one wishes to study, investigate, or interpret. In general, political research aims to understand the links between the dependent variable (the outcome of the process one wants to explain), the independent variable (the structure, the context, the cognitive frame within which this process unfolds), and the intervening variable (the factor that exerts an influence on that given process). As Hans Keman observed, the question to understand is which independent variables can account for the variation of the dependent variable across different political systems.

Political science employs several methods—four in particular—to check its research process and falsify its results: the experimental method, the statistical method, the case study, and the comparative method.

The Experimental Method

The experimental method can rarely be employed by political scientists. In contrast to the natural scientist, the political scientist cannot hope to study politics in a laboratory, in which the intervening factors can be kept constant so that the causal effects of the independent variable on the dependent variable can be reconstructed. Politics is rather more complex than what is studied in a science laboratory, above all because it involves factors that cannot be isolated and is structured by interactions that cannot be separated. Moreover, whereas in natural science the objects of study may be inanimate, the same cannot be said of politics and of those who engage in it. Politics is activated by actors (heads of governments, ministers, members of parliament, party leaders and activists, members of movements and associations, and citizens) who continually learn from their experience, thus modifying their behavior, even in the absence of variation in the independent variables.

The Statistical Method

The statistical method, however, is ever more widely employed by political scientists. It obviously presupposes the availability of numerical data. Those data are the product of a standardized process of quantitative measurement of aspects of political life—standardized because the same criteria of measurement can (and should) be used in different contexts. It goes without saying that, to be effective, this method requires a large number of quantitative and reliable data. However, whereas the experimental method presupposes the existence of a cause-and-effect relation, the statistical method does not. It provides information, but as such, it does not suggest an interpretation. Thus, without the support of a theory concerning relations between variables (measured quantitatively), the statistical method is destined to be of little use in scientific research. Moreover, when using this method, one is forced to neglect factors that are not easily quantifiable, such as cultural factors.

The Case Study Method

The analysis of the idiosyncratic factors mentioned above, instead, constitutes the object of the third method at the disposal of political research: the case study. This method has also characterized historical research and is designed to collect the largest possible amount of information (quantitative and qualitative) on a specific country (or other political entity). This method is also referred to as ideographic. A case study may have a purely descriptive purpose or, instead, may have an interpretative goal or even be designed to generate hypotheses (if not theories) susceptible to generalization (e.g., Alexis de Tocqueville's study of U.S. democracy, from which several hypotheses as well as a theory on the tendencies of Western democracies have been derived). This method is largely used in the United States, as is evident even from the comparative politics section of the book reviews in the American Political Science Association journal *Perspectives on Politics*, where there is an abundance of studies on single countries. This is not so (or rather less so) in Europe, where comparison generally involves the study of several cases.

The Comparative Method

There can be no doubt that the main method at the disposal of political science is the comparative method. In a now-classic 1970 essay, Giovanni Sartori pointed out that all political science presupposes, even if implicitly, a comparative frame of reference. The same author would write repeatedly that he who knows only one thing knows nothing at all. The comparative method requires, first, that the object to be compared be defined, next that the units used to compare and the time period to which the comparison refers be identified, and, finally, that the properties of those units be specified. Whereas the statistical method presupposes that the variables of many cases can be quantified, the comparative method is generally applied to a more limited number of cases. These cases may be very similar—that is, the strategy of the most similar research design—or they may be very dissimilar—the strategy of the most dissimilar research design. The first strategy allows more in-depth comparisons, whereas the second yields broader comparisons. The choice of strategy depends on the purpose of the study (see Table 1).

Comparative politics has become increasingly identified with the comparative method, to the extent that that method has become the defining characteristic of the academic discipline. It has permitted the construction of a particular type of scientific explanation based on correlations, as it assesses the validity of an explanation by assuming a correspondence between the properties of the independent variables and those of the dependent variables. But how and where do such correlations operate? The comparative method does not provide a clear answer. Regarding the *how*, the correlations operate if there are actors that activate them. How then should political agency be conceptualized? There is no univocal answer to this question. Regarding the *where*, the comparative method has to rely on few cases. The method of correlations has proved to be satisfactory when

Table 1 Methods of Comparative Analysis

Method	*Strength*	*Weakness*
Experimental method	(Few applications)	(Few applications)
Statistical method	Extensive information	Little interpretation
Case study	In-depth knowledge	Generic comparison
Comparative method	Hypothesis testing	Limited application

those cases are not only limited but also homogeneous (e.g., established democracies, post-Soviet democratizing countries). But if one wishes to broaden the perspective by comparing many nonhomogeneous cases (e.g., democracies, nondemocracies, and emerging democracies), what happens to a method based on correlations between uniform variables? It is no coincidence that political scientists have "regional" knowledge that cannot be easily applied in other "regions."

Conclusions

Certainly the buoyant growth of available data on the various political systems of the world has permitted complementing the statistical method with the comparative one, thus arriving at some classificatory systems that manage to comprise more cases belonging to different "regions," thereby making the analyses less Western centric. Yet the problem raised years ago by Sartori remains unsolved. Sartori argued that empirical concepts are subject to a sort of trade-off between their extension and their intensity. If a concept is applied to a large number of cases (extension), it will display only a limited ability to generate valid explanations for each case (intensity), and vice versa. When concepts are so abstract that they can be applied to the entire world, then, analytical vagueness is inevitable.

Comparison is a method used to test research hypotheses. The comparative method is generally employed by scholars of comparative politics, although it is increasingly complemented by the statistical method. The end of the Cold War in the early 1990s has brought about an enormous growth in the number of formally sovereign countries. Simultaneously, the available information resources (e.g., databases) also grew significantly, thus making it possible to have recourse to statistical comparisons. Yet a method is a tool at the disposal of researchers who take a scientific approach. It is a way of organizing the conceptual relations between the factors that are thought to structure the problem under investigation. Accordingly, there is a need for a theory that identifies those relations on the basis of a logically consistent argument. The theory informs us what to research and why. Bereft of a theoretical orientation, even the most sophisticated method provides only information or description. How to use or interpret the data depends on the theory.

Theories of Comparative Politics

There are many theories in comparative politics. The main theories generally have a focus on institutions; they are variations of the institutionalist approach. Institutionalism not only constitutes the main branch of the theories of comparative politics but also stands at the origin of political science as a whole. Without harking back to Aristotle, the genealogical tree of comparative politics had its roots at the beginning of the 20th century in the sociological constitutionalism of scholars such as Max Weber and Vilfredo Pareto and the legal constitutionalism of scholars such as James Bryce and Woodrow Wilson. From there developed, after World War II, the comparative historical sociology of scholars such as Stein Rokkan and Harry Eckstein as well as the comparative political science of scholars such as Robert A. Dahl, Samuel P. Huntington, and Giovanni Sartori. These roots obviously have not prevented the subsequent emergence of noninstitutionalist developments, frequently deriving from political scientists' use of theoretical constructs from other social sciences. Examples include behavioralism (derived from social psychology), structural functionalism (derived from anthropology and subsequently from sociology), and systems theory (derived from the new cybernetic sciences).

All these new developments had to live alongside the institutionalist approach (or "old" institutionalism), which never ceased to exert its influence on the comparative political research of the post–World War II period. Since the 1980s, the old institutionalism has been superseded by theoretical developments that have merged in the new institutionalism—new because it is distinct from its predecessor owing to its nonformalistic vision of institutions and norms. As noted by Ira Katznelson and Helen V. Milner, there are significant differences within the new or neoinstitutionalism. Some are microlevel institutionalist theories, as they take their point of departure from the preferences or interests of individual political actors or collective political actors understood as unitary. Others, instead, are macrolevel institutionalist theories, as they depart from supraindividual aggregates. Still others are mesolevel institutionalist theories, as they study the cognitive interactions between determined (but not limited) institutional configurations as well as the actors that operate within them. Whereas the macrolevel and mesolevel approaches treat the institutions (i.e., their structural characteristics or their cultural codes) as the independent variable, in the microlevel theories, this role is played instead by the preferences or the interests of the individual actors.

Rational Choice Institutionalism

The Theoretical Apparatus

The main microlevel theory of comparative politics is rational choice institutionalism. It constitutes an adaptation to political science of theories developed in the field of economics. The scholars who have most influenced the rational choice theorists are Anthony Downs, William Riker, Mancur Olson, and Douglass North. The point of reference for rational choice theorists is the neoclassical economic model, which conceptualizes collective action as the outcome of the behavior of individuals aiming to maximize their own utility (i.e., their own egoistic interests). Similarly, the political theory of rational choice postulates that the individuals who participate in politics are rational actors, acting on the basis of strategic considerations in order to maximize their utility. The actors' preferences are formed outside the political process. They are exogenous with respect to the interaction between the actors. Accordingly, individual preferences can be treated as the independent variable in this approach.

The basic unit of analysis of the rationalist theories is the actor. Politics is a game between individual actors or between collective actors understood as unitary subjects. As Riker has argued in many writings, the aim of this theory is to explain how collective action emerges in a multiactor game and to examine the microlevel foundations of processes that give rise to macrolevel effects. This does not imply that the macrolevel effects are necessarily rational (on the social level). Rather, as pointed out by Margaret Levi, a scholar working within this approach, collective action may be irrational even though individuals act in a rational way, unless they are subjected to the constraints of specific rules in the pursuit of their interests.

Institutions matter because they make collective action possible by lowering the transaction costs between the actors, by furnishing reliable information on the rules of the transaction itself, and by sanctioning free riding, thus making individual behavior predictable. Rational choice institutionalism assumes that institutions are necessary because they make possible virtuous interaction (i.e., cooperation) between the actors. Without the rules guaranteed and promoted by institutions, the game would become uncooperative. Institutional equilibrium obtains when none of the actors has an incentive to question the status quo because no actor is able to establish whether a more satisfactory equilibrium may result from doing so. Such equilibriums are defined as Pareto optimal, with reference to the well-known Italian economist and sociologist Vilfredo Pareto (1848–1923). For rational choice theory, as it is employed in comparative politics, institutions and norms are intervening variables—factors that intervene from the outside to regulate the interactions between individual actors.

Rational choice institutionalism is a testable theory because, once the institutional framework that regulates behavior is known, it can yield falsifiable statements. It establishes a correlation of events identifying the causal mechanism that links the independent with the dependent variables. By assuming that individual behavior (as well as the behavior of collective actors whose actions are taken to be unitary), in any context whatsoever, is driven by the maximization of individual utility, rational choice theory is able to claim that its analytical framework has universal validity. Moreover, this scientific claim is supported by quantifying variables so that the rational choice theory can be expressed in a formal mathematical language. Indeed, according to these theorists, political science may legitimately compete with economics as long as it adopts the highly formalized analytical apparatus characteristic of the latter. Hence, rational choice theory derives its extraordinary academic and scientific success, especially among the community of U.S. political scientists, where it still sustains a strong hegemony of neoclassical economics within the social sciences.

Criticisms

Rational choice theory has not been spared criticism. First, it has been pointed out that an analysis that assumes institutions to be ahistorical entities ends up being overly abstract. Institutions are not aseptic rules but structures that are permeable to history. The same institutional structure may produce different effects in different historical periods. Second, it has also been shown that institutions constitute coherent and interconnected agglomerates of regulative structures. As a result, it is hardly plausible to examine a given institutional variable in isolation, assuming that the other institutional variables will remain constant. In reality, there is a reciprocal interaction between them, since the institutional variable considered may produce different effects if the variables with which it is linked are subjected to specific and unforeseen influences. Third, it has been argued that rational choice institutionalism runs the risk of turning into a remodeled functionalist theory, to the extent that it explains the existence of an institution with reference to the effects that it produces.

Fourth, it has been shown how an intentionalist conception of human agency lurks behind the theory of rational choice because it assumes that the process of creating an institution is intentional—that is, controlled by actors who correctly perceive the effects of the institutions created by them. Indeed, it is rarely the case that the actors intend to create institutions or that the actors correctly foresee the future impact of institutions. Moreover, rational choice institutionalists conceive the formation of institutions as a quasi-contractual process marked by voluntary agreements between relatively equal and independent actors, almost as if they still found themselves in a state of nature. Again, this is rarely the case. The elegance of the formal theory is insufficient to compensate for the weakness of the hypotheses of rational choice institutionalism concerning the driving forces of individual choice and the persistence of institutions.

Rational choice has been shown to be convincing in the analysis of microphenomena (e.g., why a certain decision was made in a committee of the U.S. Congress or why the negotiations within the intergovernmental conference of the European heads of state and government led to a certain outcome), but it is less successful in analyzing macrolevel processes (e.g., why the U.S. Congress has established itself as the most powerful legislature of the democratic world or why the process of European integration has given rise to a political system unforeseen by the states that had launched this process). Rational choice theory is generally of little use when the number of actors involved in the phenomenon under examination is large and when the timeframe is extended. It is certainly true that rational choice theory has positively contributed to raising the degree of formalization of political science. Yet as Barbara Geddes points out, it is equally beyond doubt that rational choice scholars have focused only on those phenomena that would allow such formalization. As a result, rational choice scholars have ended up studying a problem not because of its political relevance but of its ability to be formalized.

Historical Institutionalism

The Theoretical Apparatus

The main macrolevel theory of comparative politics is historical institutionalism. It has been developed by scholars such as Paul Pierson, Theda Skocpol, and Kathleen Thelen among others, elaborating the rich tradition of the historical social sciences of the 1950s and 1960s, represented by Barrington Moore Jr., Reinhard Bendix, and Seymour M. Lipset. It differs from rational choice institutionalism to the extent that it understands institutions not simply as arrangements that serve to regulate an interactive game but as historical structures that have origins and develop independently from those that operate within them. Moreover, whereas the analytical focus of rational choice institutionalism is on the actor, the analytical focus of historical institutionalism lies instead on institutional structures and their evolution over time.

Historical institutionalists analyze institutional and organizational configurations rather than single institutions in isolation, and they pay attention to processes of long duration. They show how general contexts and interactor processes give shape to the units that organize the political process. For them, time is a crucial intervening variable in explanations of specific outcomes. The aim of the analysis is to establish the sequences and the variations of scale and time that characterize a given political process. One of the fundamental concepts of historical institutionalism is path dependence. The theory of path dependence argues that, in politics, decisions made at time t will tend to shape the decisions made at time $t + 1$. Once a given institution has asserted itself, it tends to reproduce over time. Contrary to what occurs in the economy, however, in politics, the marginal productivity of an institution increases over time, as Paul Pierson has shown.

To the individualistic outlook of rational choice theory, historical institutionalism has placed in opposition a vision of the political process as structured by institutions that have consolidated over time and thus shape this process. Also, windows of opportunity for institutional change open up under conditions of institutional crisis, but the actors, nevertheless, are constrained to act within the bounds inherited from the previous arrangements. Here, there is no heroic vision of actors as in the voluntaristic vision of agency that rational choice institutionalism assumes. Simultaneously, regarding historical macro-analyses, historical institutionalism has steered clear of the pitfalls of teleology that have frequently imprisoned historical research.

Criticisms

This theory too has been subject to criticism. First, historical institutionalism has paid less attention than rational choice institutionalism to the role of individual actors, and, in general, it has been more concerned with the possibility of agency in the historical evolution of a given institutional structure. The absence of a theory of agency has led historical institutionalists to emphasize the inertia of institutions (conceived as "sticky" structures), even though they were frequently constrained to modify or adapt themselves. For this reason, they are unable to provide precise indications concerning the chains of cause and effect that operate between the institutional macrostructures and the microlevel individual decisions. Second, historical institutionalism does not have at its disposal an analytical device to falsify its conclusions. Once a particular institution or policy has been reconstructed, it may prove difficult to imagine an alternative sequence. If every historical development is unique and if it is not possible to employ counterfactual hypotheses, then the possibility that the theory will fall into some kind of determinism is high.

Third, historical institutionalists have compared only a limited number of cases (and it could not have been otherwise, given their need to examine each case in depth), so how can valid knowledge be generated from these few cases unless supported by an extensive verification of the postulated hypotheses? It is no coincidence that the scholars who most extensively employ statistical methods criticize historical institutionalists for selecting their case studies according to the dependent variable that they wish to explain. As Gary King, Robert Keohane, and Sidney Verba have pointed out, historical institutionalism, like rational choice institutionalism, tends to select cases that fit with the approach adopted while ignoring those that do not. For example, historical institutionalists compare countries in which a political revolution occurred but do not explain why this did not happen in other countries with similar economic, social, and cultural characteristics.

Sociological Institutionalism

The Return of Political Culture

The main mesolevel theory of comparative politics is sociological institutionalism. Among the representatives of this approach, James G. March and Johan P. Olsen should be mentioned. According to sociological institutionalism, institutions are not just the source of rules for solving the problem of collective action or path-dependent structures that condition future decisions but configurations of meanings that the actors come to adopt. They solicit the formation of mental maps concerning the appropriate political behavior that guide the actors operating within institutions. They are sources of meaning independent of the actors that embrace them. Sociological institutionalism is a mesolevel theory that conceptualizes interactions between actors within institutional systems that are not limited (unlike limited systems, e.g., the committees of the U.S. Congress) and that are not extended (unlike systems such as welfare states).

Sociological institutionalism exhibits a strategic interest in culture (i.e., in meanings, symbols, common sense, ways of thinking, and cognitive frames). Indeed, it is a development of the rich tradition of cultural theory of comparative politics, which was rather relevant in political science until the 1970s. After a phase of decline, during the 1990s, the interest in political culture has clamorously returned to the stage of scientific debate. In particular, with the development of large-scale researches on social capital, the interest in political culture was and continues to be shared not only by political scientists but also by sociologists, economists, and anthropologists. Robert Putnam has been a pioneer in this regard, first conducting extensive empirical research on the relations between the social capital of the various regions of Italy and their institutional performance (Putnam, 1993) and then investigating the decline of social capital in the United States and its effects on the associationalism of that country (Putnam, 2000). In both cases, Putnam has shown how the quality of civic life forms the basis for the development of effective institutions in the context of a democratic society. This model was thus tested in a comparative research effort, directed by Putnam and published in 2002, which analyzed the relation between social capital and institutional performance in the main contemporary advanced democracies.

According to these theorists, political culture has to be treated as the independent variable and institutional performance as the dependent variable. Societies differ because they are characterized by different politico-cultural attitudes. Those cultural differences are relatively durable, even if they are not immutable. They can explain, for example, why some countries enjoy a stable democracy whereas others do not. The theorists of political culture have sought to provide an intersubjective analysis of politics, in which the intersubjectivity is conditioned (if not determined) by the cultural context in which it develops. This approach has been of considerable relevance for the contemporary world as it has provided an explanation of why institutional innovations encounter difficulties in unchanged cultural contexts and why processes of democratization find it difficult to advance in certain countries where the cultural assumptions of the previous regime have not been questioned. Yet, according to the critics, the concept of political culture risks being tautological in character. Citizens behave in a certain way because of the presence of a particular culture, but that particular culture is defined by the fact that citizens behave in a certain way. As a consequence, this concept makes it hard to explain the changes in social behavior and individual beliefs, changes that nevertheless occur regularly in contemporary political systems.

Sociological institutionalists have sought to come to terms with that debate by reducing the concept of political culture to the institutions that shape it and to the individuals that legitimate them. According to these scholars, institutions are not only rules of the game or crystallized historical arrangements but also encompass symbolic systems, cognitive maps, and moral frameworks of reference that represent the meanings that guide behavior. Accordingly, culture is transmitted by means of institutions, and because of this, they are simultaneously formative and constraining. Culture functions as the independent variable to the extent that it is institutionalized. Sociological institutionalism has undertaken a holistic analysis that seeks to recombine agency and structure in a comprehensive scheme of meanings and ways of thinking. As has been noted, anyone who has spent time waiting at a traffic light when nobody else was around will be able to understand the importance of internalized ways of thinking. In stable institutional contexts, as March and Olsen have noted, behavior is rarely motivated by the logic of utility but rather by the logic of appropriateness. Individuals behave in a certain way because that is what is expected of them in that particular institutional context. A given political culture is a component of the process that produces it, thus making the relation between the independent and the dependent variables in this process much more interactive.

Criticisms

Sociological institutionalism has not been immune to criticism. Two points are of particular relevance. First, as Peter A. Hall and Rosemary C. R. Taylor (1996) have argued, sociological institutionalism is based on the assumption of stable and legitimate institutional conditions. But what happens if one wants to explain the emergence or construction of new institutions? And what is the appropriate behavior in a situation of change? Processes of institution building inevitably tend to involve actors with diverse and contradictory cognitive frameworks and with conflicting interests. Second, the holistic approach of sociological institutionalism does not allow for a conceptualization, and even less measurement, of the contrasts between cognitive frameworks—that is, of competition or conflict between actors over how to understand and apply the predominant cultural frames. Institutions and actors stand in a relation of reciprocal influence: Whereas institutions contribute to defining the identity of the actors, sociological institutionalism must also hold that the characteristics of the latter (their decisions, their strategies, their visions) tend to define the identity of the former. Sociological institutionalism, thus, underestimates the impact of competition and conflict between actors (and between their contrasting interpretations of the appropriate behavior) on the definition of those same institutions. This implies that even consolidated institutions are cognitively less determined than sociological institutionalism assumes. For a summary of the discussion on the three institutionalisms, see Table 2.

Conclusion

It is difficult to find a theory of comparative politics that does not refer, in one way or another, to institutions. For comparative politics scholars, institutions matter. Yet these theories differ significantly with respect to (a) what is understood to be an institution, (b) how institutions are created, (c) why and when they are important, and (d) how institutions change. The different viewpoints of institutionalist scholars have also highlighted the existence of different research programs. The rationalist approach is engaged in a formidable undertaking of simplification of comparative politics, as these scholars aim to provide their research program with a microeconomic foundation. Historical and sociological institutionalists, instead, seem to be engaged in an equally formidable enterprise of complexification of comparative politics, as they start from less limited and less restrictive assumptions. The former seek to construct a theory on the basis of the actor, whereas the latter start from the structures or the meanings embedded in them.

This division ultimately refers to the problem of what is to be understood as a theory in political science (Charles Ragin, 1994). According to the rationalists, political science needs to equip itself with an epistemological structure similar to that of the most formalized social science—namely, economics. For nonrational approaches, its social utility is justified by the ability to provide conceptual frameworks within which problems of public relevance can be examined but without any pretense of becoming a positive science. The first wants to predict what will happen; the other two want to explain what has happened. Which one of the available theories to employ should be decided by the problem under investigation and not as a matter of principle. Because it is the task of political science to investigate different problems, then, as Donatella Della Porta and Michael Keating submit, the lively theoretical pluralism of the discipline should be welcomed.

Table 2 Neo-Institutionalist Theories

Theories	*Level of Analysis*	*Independent Variable*	*Intervening Variable*	*Dependent Variable (Examples)*	*Role of Institutions*
Rational choice institutionalism	Micro	Individual preferences	Institutions	• Decisions • Allocation of resources and posts	Make collective action possible
Historical institutionalism	Macro	Institutional arrangements	Time	• Systems of regulation • Public policies	Structure the context of the political process
Sociological institutionalism	Meso	Institutionalized cultural systems	Meanings	• Decisions • Strategies	Shape the identity of the actors

Issues of Comparative Politics

There are significant regional differences in the issues investigated by scholars in Africa, Asia, Latin America, Europe, and the United States since the end of the Cold War. However, those issues generally deal with the implications, functioning, and transformation of democracy. This is perhaps why institutional theories of comparative politics have become so successful. Democracy is a political regime that requires specific institutions, although those institutions may function properly if legitimated by coherent values (or political culture) diffused among the citizens. Since the end of the Cold War, democracy has become the only legitimate game in town. In 1900, only 10 countries were considered democracies, but in 1975, there were 30 such countries. In 2010, 115 out of 194 countries recognized by the United Nations were considered electoral democracies by the international nongovernmental organization Freedom House. This spectacular diffusion of democracy has inevitably attracted the interest of comparative politics scholars. Certainly, other topics unrelated to democracy have been investigated in a comparative perspective (e.g., revolutions, civil wars, ethnic strife, Islamic regimes). Nevertheless, the issues connected to democracy have represented the operational link between Western and non-Western scholars. The following sections identify a few areas of investigation within the vast literature.

Democracy and Supranational Developments

The debate on democratic models has continued to be at the center of comparative politics in the Western world. Thanks to the pioneering work of Arend Lijphart (1999), different patterns of democratic organization and functioning have been detected within the family of stable democratic countries. According to Lijphart, democracies might be classified according to the two ideal types—majoritarian democracies and consensual democracies—as a consequence of the structure of their social cleavages and institutional rules. This classification has been very important for freeing the analysis from the old normative argument, which assumed that there were more developed democracies (of course, the Anglo-American ones) and less developed democracies (of course, the continental European ones). Lijphart's classification has been revised by several authors. For some of them, such as Sergio Fabbrini (2008), the distinction between patterns of democracy concerns more their functional logic than their specific institutional properties. What matters is the fact that certain democracies function through an alternation in government of opposite political options, whereas others function through aggregation in government of all the main political options. Indeed, alternation in government takes place regularly in democracies that do not adopt a majoritarian (Westminster) first-past-the-post electoral system, such as Spain, Greece, or Germany. These democracies are competitive, notwithstanding their nonmajoritarian electoral systems. The consolidation of democracy in Eastern Europe, Latin America, and Asia has increased the number of countries to be considered for the identification of democratic patterns. In dealing with this process, Lijphart (in his subsequent works) has gradually (and surprisingly) shifted to a more normative approach, arguing that the consensual model represents a better model for the new democracies to adopt.

At the same time, the end of the Cold War and the prospect of the political reaggregation of the continent have accelerated the process of European integration; in the period from the Maastricht Treaty (1992) to the Lisbon Treaty (2009), Europe has become the European Union (or EU). The process of European integration was traditionally considered a unique experiment in international relations. International relations scholars were interested in explaining the process of integration rather than its outcome (i.e., the community system and its institutional characteristics). Since the Treaty of Maastricht (1992), a new generation of studies has started to investigate the EU as a political system. However, although the EU could no longer be considered an international regime, it could not be compared with other domestic political systems. In some cases, its supranational character has come to be considered exceptional, sui generis, unique by several observers and scholars. Or, if compared, the EU has been compared on the basis of generic or behavioral criteria, such as the style of decision making, ways in which political leaders interact, attitudes in managing public policies, and relations between interest groups. Lijphart (in his book of 1999) considered the EU as a case of consensual democracy.

A different comparative approach has been taken by other authors. Based on specific institutional criteria, Fabbrini (2010) has argued that the EU is a political system organized around multiple separations of powers. In the EU, there is no government as such, as in the parliamentary or semipresidential systems of its member states that are organized according to the principle of the fusion of powers. Contrary to systems of fusion of powers, the system of multiple separations of power functions without a government as the final locus of decision-making power. Such systems are proper unions of states rather than nation-states—in particular, unions of asymmetrically correlated states. Because of this (structural and cultural) asymmetry, such unions cannot accommodate the centralization of decision-making power. If institutions matter, then to classify the EU as a consensual democracy appears highly unconvincing. Like other democratic unions of states, such the United States and Switzerland, the EU is a species of a different democratic genus, and could be called a compound democracy. Asymmetrical unions of states can be subsumed neither under the model of consensual democracy nor under the models of majoritarian/competitive democracy, because they have neither a government nor an opposition. One might argue that they are Madisonian systems functioning on the basis of checks and balances between institutions and not between political options as in fusion-of-powers democracies. The classification of democratic patterns, if it is to take into consideration institutional systems, needs be enlarged to a more comprehensive typology. The development of the EU has allowed comparative politics to overcome national borders and apply its tools, concepts, methods, and theories to the study of a supranational political system. At the same time, the EU has also been compared with other regional organizations, a comparison that has shown the difference between political and economic regionalism. The comparative analysis of politics has been relaunched by the development of the EU.

Democratization and Consolidation

The end of the Cold War also had dramatic effects in the non-Western world, ushering in a new democratic era in Africa, Latin America, and Asia. In these regions, a process of democratization started with the aim of creating regimes able to provide, as it was claimed in many quarters, security in the sense of protection against widespread and arbitrary violations of civil liberties. Many political elites of new democratizing countries seemed to share the belief that a democratic regime has an intrinsic value. This belief was epitomized by the introduction of democratic rule in South Africa in 1994. On these empirical bases, the 1990s registered a diffusion of studies on democratization (e.g., Geddes, 2007), studies that benefitted largely from the previous generation of works on the democratization of Southern Europe and Latin America, such as that by Leonardo Morlino (1998).

Democratization has entailed the introduction of reforms aimed at limiting the role of the state in the political sphere. It has been about restoring political pluralism, whereby different political and civic organizations participate in the political process without hindrance. According to Ben O. Nwabueze (1993), democratization was meant to enhance transparency and accountability. Although most countries in Africa and Asia have introduced political changes in their polities, the crucial question, especially in Africa, has continued to be whether democratization is reversible or not. As Larry Diamond and Marc Plattner, among others, have pointed out, in some cases, newly established democratic orders have devolved into pseudodemocracies. Despite this, Africa, Asia, and Latin America have made tremendous progress toward democratization, although North Africa and the Middle East have yet to make a major step in this regard. Especially in the 1990s, scholars of comparative politics devoted their work to seeking to explain why some countries and not others were successful in transitioning from nondemocratic to democratic systems. Some studies focused on the strategic role played by individual leaders (Nelson Mandela in South Africa or Mahathir Mohamad in Malaysia) as key drivers of change and guarantors of political transformation. Others emphasized the role of specific institutional settings for supporting the democratization of a country. Still others have investigated the role of civil society in fostering or contrasting democratization, according to an approach not dissimilar to the social capital approach.

With the diffusion of democracy in Africa, Asia, and Latin America in the first decade of the

21st century, scholars have started to investigate the conditions that have helped or impeded the consolidation of democracy after two decades of repeated elections and broad institutional reforms. The challenges of democratic consolidation have been greater in Africa than in other parts of the world. The democratization of Eastern European and Latin American countries has been largely supported by regional organizations (e.g., as the EU or Mercosur [Mercado Común del Sur]). One of the conditions for participating in these organizations and for enjoying their economic benefits is that democratic principles should be respected. This has not been the case in Africa because of the fragility of the African Union. Although there are still cases of conflict and instability in countries such as Venezuela, Peru, Timor-Leste (East Timor), Thailand, Georgia, and Kyrgyzstan, these are isolated cases. It is in Africa that the consolidation of democracy has continued to be an open question. Although Africa has made a big step toward democratization, democracy is far from having been consolidated. Democratic reversal has continued to be a likely possibility for most African countries. The political crises in Kenya, Zimbabwe, and Madagascar in the past few years are cases in point. Democratic consolidation requires, among other things, a government turnover, and most countries in Africa are yet to undergo this crucial test. Even in countries that have maintained stability, such as Botswana and Namibia, one-party dominance remains a key feature of their democracy.

Quality of Democracy and Development

Following pioneering works on the quality of democracy, such as that by Larry Diamond and Leonardo Morlino (2005), scholars have investigated the same in the countries of Africa and Asia, reaching very critical conclusions. In Africa, due mainly to the low levels of development and widespread poverty, the traditional debate on democracy versus economic development has been very much alive (Adam Przeworski, Michael Alvarez, José Antonio Cheibub, & Fernando Limongi, 2000). In the 1990s, the return to multiparty democratic systems in African countries raised hopes and expectations for development. After almost two decades of multiparty elections, however, those hopes have begun to dwindle as development is slow and poverty levels have remained high. A number of surveys, such as those by the Afro Barometer Group, United Nations Economic Commission for Africa (UNECA), and Transparency International and several others based on regional and individual country comparisons, have been conducted in Africa on the subject of public perceptions of good governance, the effective delivery of services by the state, and levels of corruption and performance on human development indicators. In Asia, similar studies on public perceptions of democracy and the effects of government performance on development have been common.

Investigating developments in Africa and Asia, comparative politics scholars have come up with indicators that purport to measure democracy and good governance. The World Bank has been a leading institution in asserting that good governance is the basis for economic success. It has argued that those countries that have successfully instituted rule of law, established a culture of regular free and fair elections, and minimized corruption have been able to attract foreign private investors and thereby performed much better in development than those that have been unable to do so. Other organizations based mainly in Europe and North America, such as Freedom House, International Institute for Democracy, and the United Nations Development Program, have come up with criteria for assessing performance in political and civil rights, democracy, governance, and economic performance. In Africa, the UNECA has developed an elaborate survey methodology seeking to evaluate public perceptions of the state in various countries, comparing their performance on some key governance indicators, such as rule of law, freedom and fairness of elections, women's participation in politics, levels of corruption, and the effectiveness of the checks and balances between the core institutions of governance (with a particular focus on the independence of the judiciary).

Africa experienced poor governance and rampant corruption in the decolonization decades of the 1970s and 1980s, in part because of the diffuse corruption of public officials and governors. As a result, since the 1990s, Africa has been under pressure from international organizations and local reformers to embrace governance reforms. The African Union has acknowledged that good governance continues to be a challenge in Africa. It has therefore introduced the African Peer Review Mechanism and the African Union's Convention on Preventing and Combating Corruption and Related Offences for improving the standards of governance in the continent. Finally, some investigations have shown a correlation between human development and political stability. Indeed, human development has evaded most countries of the developing world because political instability has continued to be a major and recurrent problem. Examples include Algeria, Burundi, Kenya, Nigeria, Rwanda, Somalia, Sudan, and the Democratic Republic of the Congo in Africa; parts of India, Nepal, and Pakistan in Asia; and Colombia and Guatemala in Latin America.

Democracy and Constitutionalization

The study of constitutions and constitution-making processes features prominently in the comparative politics discourse, especially in the newly democratizing countries. The return to democratic rule has ushered in new debates on how best to craft and reform constitutions and helped ensure that constitutions facilitate democratic governance and protect human, ethnic, cultural, and other rights that characterize complex postcolonial and postconflict societies. In this regard, in Africa, the constitutions of Namibia and South Africa have been considered good examples because of their racial and ethnic balancing, power-sharing mechanisms, protection of basic human rights, and enshrined checks and balances. Constitutional reforms have been introduced in Kenya, Zimbabwe, Swaziland, and Lesotho to deal with internal conflicts. In these and other cases, the reform of the national constitution has been considered necessary for engineering peaceful political succession.

Many countries, especially in Africa, have set a two-term limit for the presidency. In practice, this innovative measure has sparked controversy and conflict in Malawi, Namibia, and Zambia, among others, where the outgoing president has sought to change the constitution to ensure that he could serve a third term. The third-term issue has thus become central in political debate. Succession is an area that has attracted interest and controversy in Africa and in some Latin American countries, such as Venezuela and Honduras. The succession not only to the state's presidency but also to the party leadership have become political issues. The cases of Botswana, Kenya, Malawi, Namibia, Nigeria, Uganda, South Africa, and Zimbabwe show that the African continent has not yet resolved these issues. Succession in the party has been a concern because it has been closely linked to succession in the presidency. As a result, political scientists such as Roger Southall and Henning Melber have focused on the role of former presidents, the legacies of political power, and the importance of term limits.

The issues of constitutions and constitution making have been central in the European debate also. Through the 1990s, all the Eastern European countries had to redefine their constitutional settings drastically. At the same time, the process of enlargement of the EU has accelerated the search for a new constitutional setting able to guarantee the functioning of a regional organization encompassing (in 2010) 27 member states and half a billion inhabitants. The first decade of the 21st century has been the constitutional decade of the EU, although this decade has witnessed the amendment of the existing treaties rather than the approval of a new and encompassing constitutional treaty. Even the contrasted process of constitutionalization of the EU has appeared less exceptional when compared with the experience of other compound democracies. This debate has also led to a vivid discussion on European citizenship and more generally on how to guarantee human rights in a multilevel supranational system.

Democracy and Representation

Political representation has changed dramatically since the end of the Cold War. In established democratic systems, starting with the United States, political parties have entered a long phase of downsizing and restructuring. Mass political parties have become icons of the past. Political parties have become state agencies in Europe (as argued by Richard Katz and Peter Mair) and electoral committees in the United States (as argued by Sandy Maisel). On both sides of the Atlantic, they have developed as supporting structures of the leader at the electoral and governmental level. A vast literature has shown how parties have been integrated by other actors in the electoral arena. While in the West scholars have been investigating the consequences on governance of the decline and transformation of political parties, in the non-Western world, the research issue has been the opposite: What role can newly founded parties play in promoting democracy? In democratizing countries, strong political parties have been considered a necessary pillar for supporting democratic consolidation (M. A. Mohamed Salih, 2003). Political parties have generally been weak and constrained by a lack of resources. This has also drawn attention to the issue of their funding. The weakness of political parties has been fed also by their tendency to fragment along ethnic lines or to rally around a founder patron who has often constrained their ability to institutionalize and practice internal democracy. Opposition parties, in particular, have been conflict ridden and fragmented throughout the new democracies, leading to their poor showing at elections.

In Africa, Asia, and Latin America, comparative politics has focused, in particular, on the functioning of elections, thus contributing to identify criteria for evaluating the legitimacy of the electoral process. The issues have concerned the impact of election rules on the legitimacy of emerging democracies, the institutional conditions that make elections free and fair, the role of electoral management bodies as institutions of governance, and the implications of observation and monitoring on the outcome of the electoral process. These studies have inspired specific regulations, such as the Southern African Development Community's (SADC) *Principles and Guidelines Governing Democratic Elections* and the Electoral Institute of Southern Africa/Electoral Commissions Forum's *Principles for Election Management, Monitoring and Observation in the SADC Region*, both made public in 2004. That notwithstanding, in many African countries, elections have not called into question the power of former liberation movements or ruling parties to dominate domestic politics.

Weakened parties in the European parliamentary democracies have contributed to the decision-making decline of legislatures, thus leaving larger room for maneuvering to the executives. Indeed, parties have come to be controlled by their leader once in government, thus justifying, as noted by Thomas Poguntke and Paul Webb, a process of presidentialization of politics in modern democracies. This process has not concerned separation of powers systems; the U.S. Congress has continued to be the most powerful legislature in the democratic world. Although for different reasons, in non-Western democracies also, parliaments have continued to remain weak institutions. For instance, it has been argued that African parliaments have been weak because of both one-party dominance of the country and an overly powerful executive, although other studies have pointed to a lack of public support for them.

To foster such support, proposals for integrating traditional institutions of leadership and governance into modern structures of representation were advanced. These traditional institutions are still influential at local and regional levels in many countries around the world. In Africa, the institution of chieftaincy, or even kingdom, is still important in people's daily existence. Political efforts to sideline, and in some countries to abolish, these traditional institutions have not succeeded. Many countries, in Africa especially, have molded their modern governance institutions around traditional ones. Botswana, Namibia, and South Africa present a model that seems to have struck a balance of mutual respect between modern and traditional institutions. In other countries such as Zimbabwe, chiefs have been directly co-opted into parliament, while in other countries such as Swaziland, it has been the traditional monarchy that has sought to incorporate modern political institutions.

The return to institutional approach in comparative politics is linked to global support for democracy. Many development agencies (e.g., as the U.S. Agency for International Development, the United Nations Development Programme, and other nongovernmental organizations) have devoted significant resources to the strengthening of the institutions of governance. In particular, across the former Soviet Union, Africa, Asia, and Latin America, the focus has been on strengthening parliamentary and judiciary institutions, and academics and practitioners are studying these programs to assess their impact. A substantial body of comparative politics literature on development assistance has thus emerged.

Comparative Policy Analysis

The interest in the organization and functioning of democratic regimes has inevitably promoted research on the latter's performance by scholars of comparative politics. Starting with the 1981 volume edited by Peter Flora and Arnold J. Heidenheimer during the 1980s, the comparative approach to policy analysis came to be adopted by many scholars, in particular for understanding the historical development of different welfare systems and for explaining the different features of welfare policies in Western Europe and the United States, as shown by the contributions of Gøsta Esping-Andersen, Paul Pierson, and Francis Castles. One might argue that the analysis of Western welfare policies constitutes the starting point of comparative policy analysis, and even today, it represents its core business.

However, with the end of the Cold War and the development of the globalization process, the comparative study of welfare systems has come to include the analysis of new social risks, risks that epitomize the pathology of contemporary postindustrial societies, as argued in the volume edited by Peter Taylor-Gooby in 2004. At the same time, the diffusion of the process of globalization has made the Western experience with welfare systems less peculiar than in the past. The problems associated with guaranteeing social security in market societies have come to be shared by many countries—developing as well as developed, democratizing as well as democratic. Market globalization has generated new social externalities, bringing to the center of comparative policy analysis new problems, such as immigration and environmental issues. In this regard, one has to consider the 1997 volume edited by Martin Janicke, Helmut Widner, and Helge Jorgens on national environmental policies and the 2007 volume by Eytan Meyers on international immigration policy.

Starting in the 1990s, the field of public policy has also seen a dramatic expansion. Globalization and Europeanization have not only created new problems, but they have pressured international institutions to promote thorough investigations of the economic, financial, administrative, and political performance of the various countries, investigations accompanied by frequent, specific recommendations for public policy. The conspicuous development of comparative policy analysis has been made possible by an easier access to empirical data concerning the various members of international and regional organizations. The latter have also supported specific policy priorities, introducing new criteria for evaluating countries' approximation to the expectations for good governance, including the call for the promotion of gender equality. Indeed, since the second half of the 1990s, gender policy has emerged as a significant policy interest, and not just in developed countries, as shown by the 1995 volume edited by Dorothy McBride Stetson and Amy G. Mazur. This policy field has been finalized not only to investigate the structural constraints on women's participation in the economic and social life of a country but also to identify resources and practices that might promote women's interests in the various national contexts.

The gender approach to public policies has thus called into question the normative premises of many public policies. In fact, the latter have been based on a culturally defined view of family organization, social needs, and individual expectations. The same concept of security has been redefined to meet new social preferences and personal attitudes. Indeed, the comparative study of public policy has led to a greater understanding of the value structures of contemporary societies, thereby helping combine empirical analysis and normative assessments.

The Future of Comparative Politics

Comparative politics (as an academic discipline) has been a major success. In the post–Cold War era, our knowledge of democratic, democratizing, and nondemocratic political systems has grown enormously. The methodologies, concepts, and theories of comparative research are widely used by political scientists. With the exception of the United States, where "American politics" scholars constitute the majority, comparative politics has become the central subdiscipline of world political science. Indeed, as noted by Henry Brady and David Collier, the discipline as a whole has been defined epistemologically by the debate in comparative politics. However, the success story of comparative politics has reached a critical juncture. Comparative politics is encountering issues with confronting the political problems of a globalized world. Its methods and theories face difficulties when applied to processes that transcend state borders and undermine the structure of the traditional political relations on which comparative politics is based.

The historical transformations that have occurred after the end of the Cold War have called into question the concept of sovereignty itself, which is the foundation of the study of comparative politics. With the processes of globalization that have unfolded tempestuously since then, the external and internal sovereignty of the nation-state (the basic unit of analysis of comparative research) has been eroded. Simultaneously, the complexity of political systems and their external relations has increased to such an unprecedented extent as to give rise to a complex interdependence. This complex interdependence is changing the nature, powers, and outlook of the units used by comparative analysis for the study of politics. It simultaneously disarticulates domestic and international politics, creating more levels of correlation between variables, levels that are not necessarily connected with each other.

This being so, it becomes increasingly less plausible to establish what constitutes an independent cause of a dependent outcome. If domestic political systems are not independent of external processes and if the actors that operate within them do not have the ability to act as agents that connect cause and effect, then the fundamental preconditions of comparative analysis are being eroded. Therefore, it is increasingly less likely to assume that the various political systems are distinct entities, because in reality, they are not.

Globalization and Europeanization

Domestic political systems are embedded in an institutionalized international context that noticeably constrains the autonomy of their decisions. The majority of nation-states are members of international economic organizations (e.g., the World Bank, the International Monetary Fund, and the World Trade Organization) or of military alliances (e.g., the North Atlantic Treaty Organization [NATO] and South East Asia Treaty Organization [SEATO]) where decisions are made (with their participation, of course) that have significant domestic implications. Over and above this, the sovereignty of the nation-state is also being eroded from below. At least in Europe, the unitary and centralized state is being superseded. New regional and municipal authorities have emerged and have become institutionalized. Not only has the room for maneuver of the national or central authorities been reduced, but the cognitive context itself, within which their actions unfold, has changed. Our conception of the state has been modified.

Globalization, as noted by Philippe Schmitter, has become the independent variable in many national contexts. It has reduced the impermeability of the nation-states to external pressures. This, in turn, has weakened relations between the citizens and the institutions of those states, as soon as the former have become aware that the latter are unable to respond to their demands. The legitimacy of public institutions has been further reduced by the growing role that noninstitutional actors have acquired in the context of globalization. These actors comprise companies, associations, and transnational nongovernmental organizations that operate outside the border of single states, and they have contributed to the emergence of new supranational regulative systems or international regimes. One may claim that no nation-state (not even the most important ones like the United States or China) is able to control domestic decision-making processes, autonomously steer its own economic dynamics, or develop its own separate cultural identity.

If globalization has challenged the assumptions of comparative politics, this is even truer in the case of the Europeanization induced by the process of European integration. There are many definitions of Europeanization. According to Vivien Schmidt, it consists in the process through which the political and economic dynamics of the EU have become part of the institutional and cultural logic of domestic politics. However one defines Europeanization, there can be no doubt that it consists of the implementation, at the level of the single member states, of institutional procedures, public policies, and cognitive frameworks to address the domestic problems deriving from the EU level. No member state of the EU (not even those most proud of their own traditional sovereignty, e.g., the United Kingdom and France, or those most proud to have recently regained sovereignty, e.g., Poland and the Czech Republic) can decide (in myriad areas of public policy) on the basis of autonomous considerations. This does not mean that the European nation-states have disappeared. Rather, it is even plausible to argue, as

Alan Milward has done, that European integration has rescued them, taking care of tasks they were no longer able to tackle. This, however, means that their sovereignty has been eroded and redefined. They are sovereign in some areas of public policy (e.g., defense and, partially, foreign affairs), whereas they are entirely bereft of sovereignty in other areas (e.g., monetary policy and, more generally, all areas related to the creation of the common market). In yet other areas (e.g., environmental and research policies), they share sovereignty.

It is clear that Europeanization constitutes a formidable challenge to the assumptions of comparative politics—namely, to conceive of the European states as sovereign and independent units that are autonomous in terms of decision making. In Europe and elsewhere, the dividing line between internal (domestic) politics and external (international) politics has shifted significantly (Robert Elgie, 1999).

Toward an International Comparative Politics

Globalization and Europeanization have brought radical transformations to the states and to the relations between them. These transformations call into question the traditional distinction between comparative politics and international relations. Regarding comparative politics, the sovereignty of nation-states has been fundamentally questioned. Sovereignty has been unpacked, fragmented, and segmented, thereby challenging a long normative (and hypocritical) tradition that assumed that sovereignty is an indivisible reality. The same supposed order of the domestic polity is dramatically belied by the fact that most of the major conflicts that occurred during the 1990s have happened within states rather than between them. Regarding international relations, it is no longer certain that the international system is the anarchic world that has formed the basic condition of interstate relations. This system is organized in several international regimes and managed by several international organizations, each of them equipped with tools and norms to peacefully adjudicate disputes between public and private actors.

Various studies have recognized that the epistemological and methodological boundaries between the two subdisciplines of political science are no longer evident. Scholars developing the foreign policy analysis approach have included the domestic structure of a given regime in their analysis, showing how it exerts a significant influence on the decisions and styles of foreign policy of a country. Scholars investigating the relations between the economic structures and the political arrangements of various countries have shown the interactions between international pressures and domestic arrangements (with their effects on the organization of markets, the construction of welfare regimes, the organization of systems of representation, and interest intermediation). One could mention the most recent literature on democratic peace, which has sought to show why democratic countries do not go to war with each other, because of the domestic constraints to which their decision makers are subjected; however, this does not imply that they are not inclined to go to war with nondemocratic countries. Finally, leading international relations scholars have continued to work with models connecting international and domestic variables.

These and other studies have called for the development of an integrated political science that is subdivided by the topics it seeks to study rather than by the units of analysis chosen (the domestic system or the international system). A political world marked by complex interdependence calls for a political science ready to experiment with new methods and new theories. A new field of study, which some scholars call International Comparative Politics, might be developed to confront the challenges of this world. However, the structure of academic careers, still rigidly organized around the distinction between the two subdisciplines, will make such development difficult.

Conclusion

The fundamental transformations induced by the processes of globalization and Europeanization have ended up questioning the methodological and theoretical self-sufficiency of comparative politics. These processes have urged scholars of comparative politics to take the international context of a country into account as an essential variable in explanations of the functioning of domestic politics. Simultaneously, the effects of domestic structures on supranational and international processes have driven international relations scholars to reexamine the methodological and theoretical self-sufficiency of their discipline. Substantial changes in the real world of politics are urging political scientists to develop methods and theories that can come to terms with the complex domestic and international forces that shape the important problems requiring study and explanation. After all, the undertaking of political science, as of all other social sciences, is justified by its ability to furnish plausible solutions to real problems. Accordingly, the profession should not be afraid to question itself, to overcome consolidated divisions between subdisciplines, and to seek new perspectives. A self-sufficient political science serves neither political scientists themselves nor the citizens of the contemporary world.

Sergio Fabbrini
Trento University
Trento, Italy

Patrick Dibere Molutsi
Tertiary Education Council
Gaborone, Botswana

See also Comparative Methods; Constitutional Engineering; Constitutionalism; Democracy, Quality; Democracy, Theories of; Democratic Consolidation; Democratization; Development, Political; European Integration; Europeanization of Policy; Institutions and Institutionalism; Policy Analysis; Policy Process, Models of; Rational Choice; Representation

Further Readings

Boix, C., & Stokes, S. (Eds.). (2007). *The Oxford handbook of comparative politics*. Oxford, UK: Oxford University Press.

Brady, H. E., & Collier, D. (Eds.). (2004). *Rethinking social inquiry: Diverse tools, shared standards*. Lanham, MD: Rowman & Littlefield.

Cain, B. E., Dalton, R. J., & Scarrow, S. E. (Eds.). (2004). *Democracy transformed? Expanding political opportunities in advanced industrial democracies*. Oxford, UK: Oxford University Press.

Caramani, D. (Ed.). (2008). *Comparative politics*. Oxford, UK: Oxford University Press.

Della Porta, D., & Keating, M. (Eds.). (2008). *Approaches and methodologies in the social sciences: A pluralist perspective*. Cambridge, UK: Cambridge University Press.

Diamond, L., & Morlino, L. (Eds.). (2005). *Assessing the quality of democracy*. Baltimore: Johns Hopkins University Press.

Elgie, R. (Ed.). (1999). *Semi-presidentialism in Europe: Comparative European politics*. Oxford, UK: Oxford University Press.

Fabbrini, S. (2008). *Politica comparata: Introduzione alle democrazie contemporanee* [Comparative politics: Introduction to contemporary democracy]. Roma-Bari, Italy: Laterza.

Fabbrini, S. (2010). *Compound democracies: Why the United States and Europe are becoming similar* (2nd rev. ed.). Oxford, UK: Oxford University Press.

Flora, P., & Heidenheimer, A. J. (Eds.). (1981). *The development of welfare states in Europe and America*. New Brunswick, NJ: Transaction Publishers.

Geddes, B. (2007). What causes democratization? In C. Boix & S. Stokes (Eds.), *The Oxford handbook of comparative politics* (pp. 317–339). Oxford, UK: Oxford University Press.

Hall, P. A., & Taylor, R. C. R. (1996). Political science and the three new institutionalisms. *Political Studies, 44*(5), 936–957.

Katznelson, I., & Milner, H. V. (Eds.). (2002). *Political science: State of the discipline*. New York: W. W. Norton.

King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.

Lijphart, A. (1999). *Patterns of democracy*. New Haven, CT: Yale University Press.

March, J. G., & Olsen, J. P. (1989). *Rediscovering institutions: The organizational bases of politics*. New York: Free Press.

Morlino, L. (1998). *Democracy between consolidation and crisis: Parties, groups, and citizens in Southern Europe*. Oxford, UK: Oxford University Press.

Nwabueze, B. O. (1993). *Democratization*. Ibadan, Nigeria: Spectrum Law.

Poguntke, T., & Webb, P. (Eds.). (2007). *The presidentialization of politics: A comparative study of modern democracies*. Oxford, UK: Oxford University Press.

Przeworski, A., Alvarez, M. E., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development: Political institutions and well-being in the world, 1950–1990*. Princeton, NJ: Princeton University Press.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

Putnam, R. D. (1993). *Making democracy work: Civic tradition in modern Italy*. Princeton, NJ: Princeton University Press.
Ragin, C. (1994). *Constructing social research*. Thousand Oaks, CA: Pine Forge Press.
Salih, M. A. (Ed.). (2003). *African political parties*. London: Pluto Press.
Sartori, G. (2009). Concept misformation in comparative politics. In D. Collier & J. Gerring (Eds.), *Concepts and method in social science: The tradition of Giovanni Sartori* (pp. 13–43). London: Routledge. (Original work published 1970, *American Political Science Review, 64,* 1033–1041)
Schmitter, P. (2009). The change and future of comparative politics. *European Political Science Review, 1*(1), 33–61.
Southall, R., & Melber, H. (Eds.). (2005). *Legacies of power: Leadership change and former presidents in African politics*. Cape Town, South Africa: HSRC Press.
Stetson, D., & Mazur, A. (Eds.). (1995). *Comparative state feminism*. Thousand Oaks, CA: Sage.

Competition, Political

Competition is understood as the struggle between two or more agents to capture scarce, limited, and valued resources within a defined system or context. It is often assumed that competition is a zero-sum game, in which one participant's gain or loss is exactly matched by the loss or gain of another. Political competition involves the struggle for power, for example, through elections in a democratic regime. It is a concept employed in virtually all subdisciplines of political science, for it is a universal aspect of human life. Although in one way or another everyone is affected by it, a clear definition of the concept remains elusive. This entry is structured as follows: First, the term *competition* is isolated from other concepts with which it has been fused or mixed. Second, using an analysis of measurement, a distinction is drawn between the two most important areas of political competition in democracies: the electoral and the governmental. Finally, the entry differentiates between the latest developments in electoral and governmental competition in new Third World democracies and those in other developed democracies.

Definition of Competition

One of the reasons for the concept's elusiveness is that the term has often been conflated with *competitiveness.* A democracy may not have competitiveness during long periods of time despite being a competitive regime (as in the example of predominant party systems). Indeed, Robert Dahl's definition of polyarchy allowed for the free exercise of political contestation in—and between—elections, but it never implied that effective competition had to occur. Thus, competition comprises a possible competitiveness but not otherwise. In the words of Giovanni Sartori (1976/2005), "*Competition is a structure*, or a rule of the game. *Competitiveness is a particular state of the game*" (p. 194, emphasis original). The literature is rather strong in claiming that the more competitiveness an election has, the higher the probability that a single vote can affect the outcome, which increases the expected utility of voting and, thereby, voter turnout. Thus, high levels of competitiveness are strongly related to the health of a democracy.

Areas of Competition: Elections and Government

Contemporary democracies are the systems in which political competition occurs par excellence. Democracy could be understood as a system of government that frames, regulates, and limits political competition in order to ensure that those elected will enjoy broader popular support and that there is equal opportunity for all competitors. Of course, not every public office to be filled in a democracy is elected (e.g., justices of the highest court) and not every aspect of democratic life is subject to political competition (e.g., respect for basic human rights is not contingent on political competition but guaranteed to all). The scope of political competition's excluded areas is still a fertile ground for theoretical discussion (e.g., why not elect high officials through lots instead of votes and why not select the best instead of the one chosen by the many?). Despite these debates about the limits of political competition in the democratic realm, political competition and competitiveness are clearly observed in two major arenas of democracies: the electoral and the governmental.

Within the realm of electoral competition, Joseph Schumpeter offered perhaps the first attempt to provide a working definition that challenged the classical, more normative doctrine of democracy. In his view, democracy was not an end in itself but rather a political competitive process or struggle for the people's vote. Schumpeter's "minimalist" definition of democracy has been a catalyst for innovation in this literature. Nonetheless, it is only after Anthony Downs that the formalization of political competition, especially among political parties, was bolstered by the analogy to markets. Political parties are crucial organizations in contemporary democracies as competition is channeled by them. The principal way to understand the competition among them was through the lens of partisanship: Parties were aggregates of vote-seeking politicians. While we usually see political competition as an ongoing and never-ending process, paradoxically, the formalization of political competition has almost the sole purpose of seeking equilibrium—that is, an outcome from which no party has an incentive to deviate—something hardly ever attained.

Downs delivered a model of political competition of candidates with respect to their ideological position in a single-issue dimension. Along a single-issue dimension, a two-party competition is expected. Parties will experience strong pressures to converge on the position of the median voter if they wish to avoid electoral defeat. Important works have been devoted to showing that this predicted convergence is just partial in real life, given that Downs's theory relies on highly restrictive assumptions; for example, much of this literature has been centered on direct democracy or voting in committees, where very few players are involved. The median voter theorem produces a Nash equilibrium, in which no player has anything to gain by changing strategies unilaterally. The literature that emanated from Downs's seminal work was strongly influenced by Duncan Black's median voter model, in probably what is one of the founding texts of social choice theory, and by Kenneth Arrow's discussion of the dilemmas of aggregating single-picked individual preferences into a collective choice. All of them were predated by Harold Hotelling's model of spatial competition—the location of different sellers in a market with respect to one another—probably the pioneering paper in public choice. Here, the market analogy is more than an analogy; this literature really does derive from the idea of market competition.

Up to that time, the literature on political competition was notably influenced, and inspired by, the bipartisan background of the U.S. political system. Although U.S. elections have tended to exhibit tremendous competitiveness, governing activity exhibited stable patterns of competition (i.e., a party in the executive and basically two large parties in the legislature of roughly the same size). The discussions arise concerning how the government performs when the party that holds the executive does not have a majority in Congress—a situation usually known as a "divided government."

Yet the Downsean perspective was not unique in observing and understanding spatial models of party competition. George Rabinowitz and Stuart Elaine Macdonald, for example, use a directional-spatial model of electoral competition in which utility is based on the intensity and direction between voters and candidates in contrast to the traditional proximity spatial model, which represents utility as a declining function of distance between voter and candidate ideal positions. The directional-spatial model of electoral competition suggests two main hypotheses: The first is that voters prefer candidates who are on their side, and the second is that given two equidistant candidates, they prefer the more intense and more credible to their cause. From the directional models, a new line of research emerged—the theory of discounting—denoting that candidates cannot fulfill all their promises, and therefore, voters discount the campaign and judge according to what they expect candidates to realistically achieve in government. This theory emphasizes the role of status quo in setting expectations about what the candidates can really do and assumes a programmatic or ideological linkage between candidates' and voters' understanding that the electorate's decisions are very focused on policy proposals or ideological stances.

An extremely lucid analysis of electoral competition was advanced by Adam Przeworski (1991). He remarked that "democratization is an act of subjecting all interests to competition, of institutionalizing uncertainty" (p. 14). Like economic competition, where there are clear limitations against a single agent imposing a price or product, political competition obeys rules among relevant actors. Thus, the institutionalization of uncertainty has three main features: *ex ante* uncertainty (anyone can win), *ex post* irreversibility (losers do not try to reverse results), and repeatability. Without any of these, democracy is hardly possible.

Measures and Theory

While there is little consensus on how to properly calculate competition—students of democracy and elections have developed different measures of this vague concept—some advances are evident in the literature. For example, it is possible to measure political competition as the frequency of alternation in power over a delimited period of time or by building multidimensional and longitudinal indices of competition for a few political units (i.e., the American states). This long-term perspective is not very useful for new democracies, in which just a few elections may have taken place. Other scholars have measured competition as the winner's percentage of the votes, the percentage margin of victory, or the raw vote margin of victory in elections. Such measures are less contingent on time and provide clearer pictures of political competition in a particular moment, yet they are heavily biased against two-party systems because margins of victory tend to be smaller in multiparty democracies.

A more useful measurement of competitiveness comes from Markku Laakso and Rein Taagepera's widely used *N* index, which is a weighted average of the number of parties that compete in a national political system, and it is expressed as $N = 1/\sum p_i^2$, where p_i is the share of seats a party has in the legislature. The *N* index is strongly related to the Herfindahl-Hirschman Index, an indicator of the amount of competition among firms in relation to the industry, which is denoted as $H = 1/\sum s_i^2$, where s_i is the market share of firm *i* in the market. Though both indices, the *N* and the *H*, were created as indices of diversity (fragmentation) of political parties or firms, they are frequently used as measurements of competitiveness. The higher the index, the more the competition. These indices also suffer from important limitations, such as other institutional features like degree of proportionality of legislatures, mathematical distributional procedures of seat allocation (e.g., D'Hondt, Hare), party registration limitations, and others.

With all their limitations, the Laakso and Taagepera or Herfindahl-Hirschman indices provide an accurate view of the correlation of power among agents in a given arena (the legislature or the market itself). Still, their major shortcoming is that little is said in regard to how the legislature or the market would perform. Knowing that counting parties is not sufficient for analyzing a party system, Giovanni Sartori in his classic work of 1976 presents a theory of the dynamics of political competitiveness in regard to whether centrifugal or centripetal forces act in a given party system, which is an improvement over the earlier ones. Recent developments have tried to make a more subtle connection between electoral results and governmental activity. For instance, David Altman and Aníbal Pérez-Liñán developed a measure that weights the sizes of the "typical" parties in government, $G = \sum g_i^2/\sum g_i$, and in opposition $O = \sum o_i^2/\sum o_i$, in the following way: $C = 1 - (|G - O|/100)$. The value of *C* tends to zero whenever the government (or the opposition) controls the whole legislature and to one if there is balance between government and opposition. This index provides a forecast of governmental activity and the relationship between the executive and legislative because it reflects the opposition's access to the legislative process rather than mere electoral outcomes.

In spite of these miscellaneous indices, a whole body of research was detached from the American model and jumped the Atlantic in search of new models of political competition, particularly in the context of governability and executive coalition theories. It is not hard to understand why this literature bloomed in the context of the study of governing coalitions in Western European parliamentary democracies. The process of forming coalitions in parliamentarian regimes involves a relatively reduced number of actors (parties) competing in a closed and rule-bound milieu (parliaments) for clear purposes (membership in the executive cabinet), where the quantity and quality of portfolios are crucial. One of the strengths of this literature is exactly that it is not circumscribed by political competition, but it necessarily touches on the flip side of the coin—political cooperation—and hence provides a more comprehensive panorama of political competition.

Thus, in the late 1950s and early 1960s, in the political competition literature, mainly via the study of coalitions, the meaning of "getting votes" shifted and came to be understood as a means to reach other objectives: office, and then policy making. Probably, the three models of party competition (votes, offices, and policies) are relevant to all parties at any instant in time, although more so at some times than at others. Doubtless, new research on governing coalitions in presidential regimes shows that parties shift their objectives in synchronization to the fixed electoral calendar. Before a national election, they are principally concerned with obtaining votes. After elections, party members strive to obtain political appointments from elected officials, and then later, policy-seeking behavior ranks first in their preferences.

This literature has developed a great deal since William Gamson's and William Riker's seminal works. Gamson theorized that parties compete for office and seek their payoff (players' utilities resulting from the outcome of a game), which is fair as it is directly proportional to the amount of resources that they contribute to a coalition. The prediction of his theory is that payoffs in any government are limited and therefore, the greater the number of parties in a coalition, the lower is the payoff for each. Resources are limited because, for example, a governing coalition cannot satisfy all policy demands simultaneously (e.g., lowering and increasing taxes). As a result, in order to not distribute unnecessary payoffs, Riker predicted the formation of winning coalitions of minimum size (meaning the number of parliamentary seats a coalition of political parties must control to receive a majority vote for investiture).

Since then, authors have devoted themselves to testing variables that presumably would affect political competition. Robert Axelrod argues that coalition parties are likely to be adjacent in a one-dimensional policy space, predicting minimal connected winning coalitions. That is, parties in an executive coalition will be adjacent to each other in a one-dimensional policy space, and also, the coalition will not include any more parties than necessary to be connected and winning. Others, such as Norman Schofield, have centered their arguments on the importance of large parties and their position in the policy space; large parties whose policy positions are located in the core of the policy space—that is, parties with policy positions such that no alternative policy is preferred by a simple majority—tend to form minority and surplus majority governments. In short, most of this literature on political competition has as its hard core the fragmentation of the party system and each party's location and coverage of the ideological space.

Although a vast number of alternative hypotheses for why and how parties compete have been put forward, three facts must be underlined. First, little research has been done on how differences in party organizations affect political competition. Most studies from a rational choice perspective treat parties as unitary actors that bargain over a set of well-defined gains—usually office or policy. There is some truth in this assumption but more error. Parties are complex organizations, where internal divisions and structures affect how they and their leaders behave. Important works in the literature have already underlined the problems attached to treating parties in this manner. Despite the fact that many of the ordinary political phenomena can be explained by treating parties as if they were unitary actors, we undoubtedly need to take into account what goes on inside parties if we want to provide an explanation of party decision making in a model of some political course of action. While some scholars stress one of these motivations, other scholars combine them in one way or another.

Second, a proper specification of political competition requires a theory whereby parties compete in multidimensional spaces, because even if one is interested only in one narrow particular policy (say taxes), the position on other issues will affect the equilibrium in that narrow particular policy. Following Downs, most commonly used theories of political competition posit a unidimensional political space where competition occurs or at least—acknowledging the multidimensionality of political life—try to isolate the particular dimension in consideration. Yet, even in the case of single-peaked preferences of political actors, the results might be unstable and inefficient as regards social decisions, as was clearly demonstrated by Condorcet's paradox of voting and Arrow's impossibility theorem. It does not mean, as Gordon Tullock shows, that unstable social decisions generate total chaos or instability of these decisions because institutions play a crucial role in suppressing the underlying unsteadiness (institutions such as agenda setters, party discipline, or procedural rules in parliaments). Some scholars (e.g., Kenneth Shepsle) call the equilibrium resulting from these regulating forces a structure-induced equilibrium.

Third, the regime divide has been minimized and brought into the discussion. Incentives for party competition are quite different in a presidential system than in a parliamentary one for four reasons. First, in a presidential regime, fewer coalition configurations are likely because one party—the president's party—is normally expected to be a coalition member. The rare exceptions to this rule so far have not undermined this normal expectation. Second, given the prominence that the president's reputation has in citizens' evaluations of governmental performance (either positive or negative), the president's coalition partners are less able to claim credit for good government performance. Thus, under presidentialism, players have fewer incentives than under parliamentarism to join or to remain in a coalition. Consequently, third, the president's approval rating has a powerful effect on parties' decisions to associate themselves with, or to distance themselves from, the government. And fourth, based on the previous statement, most of the incentives change over time in synchronization with the fixed electoral calendar. Because parties know when the next elections will take place, they have incentives to distinguish themselves from the president and start to compete in order to increase their chances of getting elected in the next election.

Despite important progress in the development of coalition theories, the jury is still out on certain issues. All theories of coalition formation and survival under parliamentary regimes make assumptions, hidden or explicit, about political parties' preferences and their concomitant logics of competition. Michael Laver provides new extensions of policy-driven political competition into a multiparty environment where voters constantly evaluate party support and switch parties to augment their expectations; parties continually readapt policy positions to the shifting affiliations of voters. Political competition is based on four clear types of party behaviors:

1. *predator* (move party policy toward the policy position of the largest party),
2. *aggregator* (adapt party policy to the ideal policy positions of party supporters),
3. *hunter* (repeat policy moves that were rewarded; otherwise make random moves), and
4. *sticker* (never change party policy).

Laver's contribution provides the first steps toward endogenizing key features of the process of competition, including the birth and death of parties, internal party decision rules, and voter ideal points.

Developments in New Third World Democracies

At the risk of making a gross generalization, some countries of what is often called the Third World are lately showing signs of new patterns of electoral and governmental competition. These patterns are characterized by the emergence of personalistic leaderships, hegemonic actors, and a tendency to institutional innovation rather than consolidation, all in the context of high levels of civic disaffection, distrust of political parties, and in general, animosity toward democratic actors and institutions. The arena of electoral competition involves new forms of citizen mobilization through a large dose of social polarization, often fed from above. Once in power, in the realm of governmental competition, leaders provide comprehensive packages of institutional reforms, endorsed by constituent assemblies and usually ratified by direct popular votes (plebiscites). The proposed reforms more often than not include the reelection of elected officials, particularly of presidents, and a concentration of power in the executive at the expense of others, primarily the judiciary. This is particularly relevant in Latin America (e.g., in Venezuela, Bolivia, and Ecuador) but is not confined to this region.

In short, political competition is probably one of the most complex concepts in political science, and as such, a clear definition of it is likely to remain hard to pin down. In any case, as a concept that touches on every aspect of political life, political competition will remain at the center stage of political science.

David Altman
Pontificia Universidad Católica de Chile
Santiago, Chile

See also Democracy, Quality; Democracy, Theories of; Democracy, Types of; Parliamentary Systems; Party System Fragmentation; Party Systems; Presidentialism

Further Readings

Arrow, K. J. (1963). *Social choice and individual values.* New Haven, CT: Yale University Press.

Downs, A. (1957). *An economic theory of democracy.* New York: Harper & Row.

Grofman, B. (1985). The neglected role of the status quo in models of issue voting. *Journal of Politics, 47,* 230–237.

Grofman, B. (2004). Downs and two-party convergence. *Annual Review of Political Science, 7,* 25–46.

Laver, M. (2005). Policy and the dynamics of political competition. *American Political Science Review, 99,* 263–281.

Przeworski, A. (1991). *Democracy and the market: Political and economic reforms in Eastern Europe.* Cambridge, UK: Cambridge University Press.

Rabinowitz, G., & Macdonald, S. E. (1989). A directional theory of issue voting. *American Political Science Review, 83,* 93–121.

Roemer, J. E. (2001). *Political competition: Theory and applications.* Cambridge, MA: Harvard University Press.

Sartori, G. (2005). *Parties and party systems: A framework for analysis.* Colchester, UK: ECPR Press.

Shepsle, K. A., & Weingast, B. R. (1981). Structure-induced equilibrium and legislative choice. *Public Choice, 37,* 503–519.

Strøm, K. (1990). A behavioral theory of competitive political parties. *American Journal of Political Science, 34,* 565–598.

Tullock, G. (1981). Why so much stability? *Public Choice, 37,* 189–202.

COMPLEXITY

The science of complexity in the natural and social sciences involves the use of formalized (mathematical or computational or both) theoretical models to study systems of interacting agents where the following hold:

1. Each agent's behavior is governed by a small set of simple rules, which often depend on local information and feedback from the agent's past behavior and from other agents' behavior.

2. Characterizing and understanding the behavior of each of the agents does not directly lead to predicting or understanding the behavior of the entire system. The local rules produce emergent patterns—stable equilibria, cycles, unstable equilibria and long transitions to new equilibria, and randomness—and emergent properties such as robustness.

3. Agents' interactions are interdependent and affect others in the system; thus, removing an agent has consequences for the system beyond merely subtracting out that single agent's direct effect on other agents with which it interacts.

This entry presents the foundations of complex system modeling and points to diverse applications in political science.

Foundations

It is said that Thomas Schelling proposed the first modern complex systems model in the social sciences. In research in the 1960s into the phenomenon of racial segregation in housing, he modeled a process where people (agents) decide where to live based on the racial mix of their neighborhood. Schelling randomly placed nickels and dimes on a checkerboard, with each coin resting on only one square and with fewer coins than available squares. He then randomly picked one of the coins, say a dime, and examined its neighborhood, meaning the squares that bordered the square on which the dime stood. There will be eight squares in the neighborhood (unless it rested on the edge of the board, which we shall ignore for this example). If the neighborhood consisted of fewer than 5 nickels, Schelling did nothing. Otherwise, he moved the dime to another place on the board where there were fewer than 5 nickels in the neighborhood. He then iteratively moved through each coin and either moved it somewhere else or kept it in place, using the 5/8ths rule for tolerance of neighbors of the other type. Schelling argued that 5/8ths was a pretty tolerant threshold, in that his coins (agents) would stay put even if half of their neighborhood consisted of the other type. He showed that even with this relatively tolerant threshold, the dimes and nickels segregated into homogeneous zones on the checkerboard. He then started over and examined what happened with other kinds of thresholds (1/2, 3/4, etc.) and how quickly, if at all, the agents segregated themselves incrementally.

Note what happens in the dynamics of the Schelling model. When a coin is moved to another area of the checkerboard, it changes its own neighborhood and the neighborhood to which it moves. A single move alters the diversity of the neighborhoods of multiple agents and thus potentially alters the decisions of multiple other agents. This is an example of the kind of feedback that occurs in complex systems. Further, it is difficult to predict simply from the decision rule of the agents how the system will behave. Schelling confessed that he could not predict in advance when and if the system would segregate into neighborhoods of all dimes and all nickels.

Today, researchers using complex systems methods typically rely on computer programs to simulate the interactions of artificial agents, examining the aggregate patterns that emerge from many instances of microlevel behaviors. (Some additional terminology: What is sometimes called agent-based modeling is the use of computation to study the links between microbehavior of agents and macropatterns in a complex system.)

As can be seen in the Schelling example, it would be misleading to say that this branch of science involves the study of highly complex models. As with any theoretical modeling in the sciences, complex systems modeling involves making simplifying assumptions to craft a model of behavior, analyzing the consequences of those assumptions, and then linking the model to real-world phenomena. The use of the word *complex* refers to the characteristics of the social system under study and not to the approach of the modeler. Many complex systems models, such as the Schelling example, are extremely simple in their construction.

Applications in Political Science

A variety of social systems have been characterized as complex, and accordingly, in social science disciplines, the use of complex systems modeling has grown. Within political science, complex systems modeling has been applied to the study of international diplomacy and war, electoral competition, voting systems, the spread of culture, the evolution of cooperative behavior, criminal behavior and punishment, the sorting of people into communities, political networks, and the development of law.

Researchers studying diplomacy and war, for instance, have devised modeling frameworks to analyze the sensitivity of international alliances to changes in resource inequalities across alliance partners. When a developing country discovers oil, what effect does this have on military or trade alliances? Complex systems models can enable researchers to analyze the system-level, often nonlinear, effects of such changes.

Using complex systems models on collective action and cooperation, researchers learned of the importance of reciprocity and forgiveness in the evolution of cooperation. For a long time, scholars puzzled over the question of how cooperative behavior arises among selfish individuals. Recently, researchers have made new discoveries about the importance of "tagging" or noninformative markers on agents that permit discrimination and discernment in choosing partners. Such tags can actually promote more cooperation among all agents rather than create groupings that shut out others. And in a practical application of these types of models, some have analyzed social systems where some agents can choose to commit crimes (citizens) and others can choose to enforce community norms (police). These models have led to discoveries about the trade-offs among different strategies for fighting crime. Societies face trade-offs between devoting resources to monitoring behavior and to stricter punishments for criminals. The consequences of particular strategies for law enforcement are difficult to understand without models that specifically incorporate feedback effects from agent behavior to the development of new agent strategies.

Researchers have also used complex systems modeling to discover the conditions under which a group of diverse individuals with different ways of framing a problem can choose more effective public policies than a group of homogeneous individuals who all frame the same problem identically. Diversity, these models tell us, can in common circumstances improve collective decision making in legislatures, councils, organizations, and committees.

Research into social and political networks has increasingly included complex systems modeling. There are models that help researchers understand better the conditions in a social network that encourage innovation and the spread of innovations among agents in that network.

Complex systems modeling is sometimes seen as a challenge to traditional game-theoretical methods of modeling. But most practitioners of complex systems modeling within political science view it as complementary to game-theoretical methods. Not only are quite a few well-known complex systems models within political science based fundamentally on game theoretical techniques, but the researchers themselves have typically published work spanning a variety of modeling tools, including game theory. Well-known complexity models depict iterated prisoner's dilemma situations. Agents in these models play repeated prisoner's dilemma games with each other, and researchers seek to discover the kinds of strategies that do well in environments with agents that can alter their strategies to get ahead.

A large portion of complexity research in political science focuses attention on models where the agents are not fully rational in the manner defined by economists and many other social scientists. That is, agents are portrayed as adaptive and boundedly rational. This certainly characterizes the Schelling agents. More generally, this means that agents display a subset of the following features:

1. they do not fully optimize their utility functions given their information;
2. they are not forward-looking in being able to predict accurately the probability of certain outcomes in the future, given their behavior; and
3. they are myopic in that they do not have information outside of some defined local "zone" of interaction or geography.

Many game-theoretical models in political science have agents with one or more of these features—modelers through various assumptions end up restricting the information or actions available to agents. Nevertheless, one generalization that can be made is that, at heart, game theory relies on the assumptions following the work of John von Neumann, Oskar Morgenstern, and John Nash that agents are rational (optimize) and intelligent (they know the game they are in as well as the modeler does). Complexity systems models can violate both of these assumptions, especially the latter. Note in the Schelling model how the agents are assumed to make decisions solely on the mix of other agents in their limited neighborhood. They do not consider the implications of their actions in the long run or what happens outside their own neighborhood (either the one they leave or the one they move to).

In fact, some commentators have identified insight gained from studying systems of adaptive, less than fully rational agents as one of the noteworthy benefits of using complexity methods. There is nothing inherent in the methodology that requires that agents be boundedly rational or adaptive, but it is clearly a common feature of such models.

Consider the recent work on political party systems using complex systems modeling of electoral competition. In the work of one researcher, adaptive political parties were assumed to be one of four types, each boundedly rational but in a different way. Hunter parties change their ideological positions in a hill-climbing pattern; that is, they keep changing in the same ideological direction as long they continue to win more votes. Predator parties change their ideological positions to mimic the most popular political party. Aggregator parties adopt the most popular ideological position among their existing supporters in the electorate. And Sticker parties do not change their ideologies. The researcher attempted to simulate the party politics of Ireland using this model, matching up the type of party to a real-world example. For instance, the researcher concluded that the Fine Gael party was like a Hunter, while the Democratic Left was like a Sticker.

What is innovative about this work is the incorporation of real-world data on micromotivations directly into the assumptions of a model. Then, after a series of complex systems simulations, the researcher compared the distribution of votes among the artificial parties and the ideological positions of the artificial parties with real-world data on parties from Ireland. The methodology permits researchers to unpack the dynamics of the system by running the simulations under various scenarios, discovering what aspects of the model are driving which system-level results.

Complex systems models are best described as a set of tools for analysis. Researchers consider what they do as providing a platform for the study of specific research questions. For example, what can cause the international system to fragment into multiple centers of power, such as what happened in the early 20th century prior to World War I? Alternatively, what causes the international system to fall into a pattern where there is one dominant, hegemonic power, such as what happened after 1990? The models enable researchers, equipped with plausible assumptions about microlevel motivations and interactions among agents, to devise counterfactuals, with careful attention to interactions, feedback, and nonlinear effects. Researchers can discover the causes of macrolevel dynamics switching from one phase (e.g., multipolarity in the international system) to another (e.g., unipolarity). In this sense, complex systems researchers provide something analogous to experimental laboratories. (There are sophisticated software programs that enable students and researchers to develop complex systems models even if they have only limited training in writing computer code.) The models are platforms for researchers to study a variety of questions in careful detail, especially about how microlevel interactions aggregate into macrolevel patterns.

Ken Kollman
University of Michigan
Ann Arbor, Michigan, United States

See also Cooperation; Game Theory; Models, Computational/Agent-Based; Nonlinear Models; Prisoners' Dilemma; Rationality, Bounded

Further Readings

Axelrod, R. (1997). *The complexity of cooperation.* Princeton, NJ: Princeton University Press.

De Marchi, S. (2005). *Computational and mathematical modeling in the social sciences*. New York: Cambridge University Press.

Epstein, J. (2005). *Generative social science.* Princeton, NJ: Princeton University Press.

Holland, J. (1998). *Emergence: From chaos to order.* Oxford, UK: Oxford University Press.

Kollman, K., Miller, J., & Page, S. (Eds.). (2003). *Computational models in political economy*. Cambridge: MIT Press.

Laver, M. (2005). Policy and the dynamics of political competition. *American Political Science Review, 99,* 263–281.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems.* Princeton, NJ: Princeton University Press.

Page, S. E. (2010). *Diversity and complexity.* Princeton, NJ: Princeton University Press.

Schelling, T. (1978). *Micromotives and macrobehavior.* New York: W. W. Norton.

Wolfram, S. (2002). *A new kind of science.* Champaign, IL: Wolfram Media.

Compliance

Compliance, simply put, is the extent to which an actor fulfils its obligations. In the realm of international relations, scholars are most concerned with state compliance and have focused their attention on the degree to which states conform to the prescriptions and proscriptions stipulated by their international commitments.

Interest in compliance has burgeoned as states increasingly construct and enter into rule-based governance arrangements to guide behavior and solve problems in a variety of issue areas, such as trade, security, human rights, and the environment. Scholarship on international cooperation, regimes, law, and institutions has focused on why states do or do not comply with their international obligations and what conditions and institutional features facilitate compliance.

The subject of compliance is often connected to the debate over the importance and efficacy of international institutions, but the relationship between compliance and effectiveness is a murky one. There is widespread agreement that compliance is not sufficient for effectiveness; however, there is a debate over whether and to what extent high levels of compliance are necessary for, or at least associated with, consequential institutions and effective agreements.

Scholars from the realist tradition are generally skeptical that international institutions and law are significant determinants of state behavior. In the view of the majority of realists, power relations are the most reliable predictor of international outcomes, and states often strive to avoid complying with inconvenient obligations. In cases where compliance rates are high, realists tend to see the outcome as the result of shallow agreements that reflect the law of the least ambitious party or as merely codifications of what states would have done even in the absence of an agreement.

Other perspectives, such as neoliberal institutionalism and social constructivism, are more sanguine regarding the potential influence of international institutions and agreements Neoliberal institutionalism, with its rationalistic metatheoretical orientation and focus on constellations of interests, sees high levels of compliance with international rules as a possibility. Moreover, compliance or even good-faith efforts that fall short of full compliance, together with other factors such as the achievement of policy goals and behavioral change, can be treated as an indicator for evaluating the impact of international institutions and agreements. This perspective holds that states operate according to a logic of consequence and that institutional design matters in influencing how states will calculate their interests and goals and evaluate the costs and benefits of compliance or noncompliance. Institutional properties and mechanisms, such as monitoring capabilities and verification systems (e.g., the International Atomic Energy Agency's safeguard system fulfills this role in the case of the nuclear nonproliferation regime), can result in increased transparency, greater information, and the creation of material (sanctions, loss of privileges) and social (reputational) costs for noncompliant behavior.

Social constructivist and normative approaches take the position that international institutions and regulations are meaningful and that their norms and rules tend to enjoy widespread compliance. As the legal scholar Louis Henkin (1979, p. 47) concluded, "Almost all nations observe almost all principles of international law and almost all of their obligations almost all of the time." Social constructivism, which has a sociological metatheoretical orientation, emphasizes the interplay between norms and identity and stresses knowledge and communication dynamics to explain state behavior. Constructivists see states as role players and have found that international institutions can play an important role in socializing states to accept and internalize particular norms and rules. In their view, decisions about compliance are driven by a logic of appropriateness that reflects intersubjective normative understandings and identity concerns. Behavior is not strictly determined by rational calculations of the state's interests but by deeper norms and shared beliefs about what actions and policies are legitimate and appropriate in international relations (Thomas Franck, 1990, pp. 205–207). Therefore, from this perspective, the decision to comply with international provisions concerning issues such as slavery, piracy, human rights, the use of chemical weapons, and the acquisition of nuclear weapons serves the important function of both shaping and reflecting a state's identity.

Promoting Compliance and Coping With Noncompliance

As Helmut Breitmeier, Oran Young, and Michael Zürn (2006) have noted, while states often comply with international rules, compliance problems remain a persistent challenge. For this reason, scholars have attempted to identify the factors that foster compliance and have debated the best way to reduce noncompliance, whether it be deliberate, inadvertent, or due to lack of capacity. The management and enforcement approaches provide contending views of the most important determinants of compliance.

The management approach has stressed the importance of transparency, dialogue among the agreement's parties, dispute resolution procedures, and technical and financial assistance to promote compliance and has downplayed the need for punitive enforcement mechanisms, such as sanctions. Advocates of the managerial school argue that this approach usefully addresses the primary types of noncompliance. The provision of resources or expertise can address noncompliance that has resulted from a lack of capacity. Clarification procedures, for example, can be used to address ambiguous behavior that other parties might perceive as possible noncompliance. Finally, mechanisms that provide transparency can deter and expose purposeful breaches. "For a party deliberately contemplating violation, the high probability of discovery reduces the expected benefits rather than increasing the costs and would thus deter violation regardless of the prospects of sanctions" (Abram Chayes & Antonia Handler Chayes, 1993, p. 18).

The enforcement approach criticizes the managerial school for underestimating the importance of sanctions and punitive measures. Enforcement scholars have argued that even in the case of environmental regimes, which are the source of many managerialist examples demonstrating high compliance levels, "enforcement plays a greater role in successes than one is led to believe and its absence is conspicuous in some notable failures" (George Downs, David Rocke, & Peter Barsoom, 1996, p. 395). They hold that in matters of deep cooperation, such as security, there is always the possibility that the benefits of cheating might be too great to be offset solely by transparency. According to advocates of the enforcement approach, elaborate verification measures attest to the importance of transparency, but to believe that the power of transparency can be exclusively depended on to compel compliance in the absence of sanctions is naive. Therefore, the enforcement approach has stressed monitoring activities in conjunction with potent sanctions. Ideally, from this point of view, multilateral agreements should be equipped with mechanisms that provide incentives (resource carrots) for compliance while applying strong sanctions (punishing sticks) for noncompliance.

Rather than treating the management and enforcement approaches as diametrically opposed propositions, some scholars have argued that well-designed international regimes or agreements would benefit from including elements recommended by both approaches. They contend that the provision of help and assistance, and clarification procedures, followed by investigation, followed by inspection, followed by detection, followed by material and social sanctions is better thought of as a continuum or process to be carried out rather than as a choice between preferred extremes. For example, a study by David Victor (1998) of the noncompliance procedure of the Montreal Protocol—the regime designed to protect the stratospheric ozone layer—found that the protocol's implementation committee was most effective when it blended the two approaches. "Management avoids the most severe and unproductive antagonism, but the credible threat of tougher actions, including sanctions, helps ensure cooperation, especially when dealing with parties who are unswayed by management alone" (Victor, 1998, p. 139).

An important study that used a relational database covering 23 international environmental regimes found elements of both the management and enforcement approaches wanting and concluded that neither was able to convincingly "explain patterns of compliance with international regimes" (Breitmeier et al., 2006, p. 110). Although this study confirmed that institutionalized enforcement measures and strong verification procedures were indeed crucial, it also found that the compliance pull exerted by legal rules that enjoyed widespread legitimacy was more important than commonly understood. According to their findings,

> adequate and even impressive rates of compliance with international environmental rules occur when appropriate incentive mechanisms are coupled with juridification, participation on the part of transnational NGOs in the rule-making process, and a responsive approach to the development of compliance mechanisms over time. (p. 112)

In the years ahead, additional work—methodological and conceptual—will be needed to help us better judge to what extent states are complying with their international obligations and more effectively evaluate to what extent international governance arrangements are actually achieving their goals. Moreover, more effort needs to be expended to improve our understanding of the relationship between compliance and the influence exerted by international institutions, regimes, and agreements. The dense thicket of long-standing and emerging transnational problems that demand international solutions ensures that the subject of international cooperation and the challenge of compliance will be of enduring relevance and interest to scholars and practitioners alike.

Charles F. Parker
Uppsala University
Uppsala, Sweden

See also Constructivism in International Relations; International Law; International Organizations; International Regimes; International Relations as a Field of Study; Liberalism in International Relations; Realism in International Relations

Further Readings

Breitmeier, H., Young, O. R., & Zürn, M. (2006). *Analyzing international environmental regimes: From case study to database*. Cambridge: MIT Press.

Chayes, A., & Chayes, A. H. (1993). On compliance. *International Organization, 47*(2), 175–205.

Chayes, A., & Chayes, A. H. (1995). *The new sovereignty: Compliance with international regulatory agreements*. Cambridge, MA: Harvard University Press.

Dorn, A. W., & Fulton, A. (1997). Securing compliance with disarmament treaties: Carrots, sticks, and the case of North Korea. *Global Governance, 3*(1), 17–40.
Downs, G. W., Rocke, D. M., & Barsoom, P. N. (1996). Is the good news about compliance good news about cooperation? *International Organization, 50*(3), 379–406.
Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International Organization, 52*(4), 887–917.
Franck, T. M. (1990). *The power of legitimacy among nations*. Oxford, UK: Oxford University Press.
Hasenclever, A., Mayer, P., & Rittberger, V. (1997). *Theories of international regimes*. Cambridge, UK: Cambridge University Press.
Henkin, L. (1979). *How nations behave: Law and foreign policy* (2nd ed.). New York: Columbia University Press.
Mitchell, R. B. (1994). Regime design matters: Intentional oil pollution and treaty compliance. *International Organization, 48*(3), 425–458.
Raustiala, K., & Slaughter, A. (2002). International law, international relations and compliance. In W. Carlsnaes, T. Risse, & B. A. Simmons (Eds.), *Handbook of international relations* (pp. 538–558). London: Sage.
Simmons, B. A. (1998). Compliance with international agreements. *Annual Review of Political Science, 1*, 75–93.
Simmons, B. A., & Martin, L. L. (2002). International organizations and institutions. In W. Carlsnaes, T. Risse, & B. Simmons (Eds.), *Handbook of international relations* (pp. 192–211). London: Sage.
Victor, D. G. (1998). The operation and effectiveness of the Montreal Protocol's non-compliance procedure. In D. Victor, K. Raustiala, & E. B. Skolnikoff (Eds.), *The implementation and effectiveness of international environmental commitments: Theory and practice* (pp. 137–176). Cambridge: MIT Press.
Young, O. (1979). *Compliance and public authority*. Baltimore: Johns Hopkins University Press.

Concept Formation

Concepts are central to the enterprise of political science. The concepts we use shape the world we see. Without solid conceptual foundations, the edifice of political science is insecure. If we fail to develop clear and precise concepts, our theoretical insights and empirical discoveries will fail to be clear and precise, too. This entry reviews major pitfalls for conceptual analysis as well as the fundamental challenges to concept formation and conceptual innovation in the study of politics.

In contemporary political science, concept formation is often regarded as a distraction, a mere prelude to serious research that is given scarce attention. Scholars sometimes ignore conceptual disputes, resolve them by fiat, or delegate their resolution to political philosophers. At the same time, a strong tradition of self-conscious and systematic concept analysis, resting on the pioneering work of Giovanni Sartori, David Collier, and others, does exist in the discipline. This entry offers an analytical synthesis that weaves together insights of conceptual debate in both philosophy and political science.

Conceptual Commitments Since its origins in ancient Greece, Western philosophy has been debating the nature and meaning of concepts. For centuries, thinkers tried to resolve one fundamental problem: the relation between the world and the mind, the objective and the subjective, things and ideas. They conceived the mind as a mirror and concepts as mental images of the outside world, as cognitive representations of objective realities that uphold the fragile correspondence between the two worlds. In the mid-20th century, the so-called linguistic turn in modern philosophy, brought about by authors like Ludwig Wittgenstein and John Austin, redefined the basic coordinates of concept analysis. It shifted the axis of conceptual debate from cognition to language and from language as a system of representative symbols ("Platonism") to language as a medium of social action ("pragmatism").

Language Acts

According to the classic conception of language, concepts are our basic units of thought. According to a pragmatic understanding of language, concepts are our basic units of (linguistic) action. In this perspective, concepts are not interior images that correspond to external realities but practical tools that allow us to do many things, many more than just putting vivid labels on inanimate objects. They allow us to threaten and promise, to bless and condemn, to give orders and to request favors, to express tenderness and anger, to know and believe, to contract marriage and christen ships, and so on. Designating objects "out there" in the external world (reference) is just one linguistic function among innumerable others.

As practitioners of social science, too, we do more than offer aseptic statements about the world. In our texts and speeches, we do more than describe and explain, more than refer to facts and associations between facts. We laud and criticize colleagues, highlight and downplay themes, support and refute arguments, persuade and dissuade readers, and so forth. However, while reference is not everything, it does play a leading role in the social sciences. All types of "speech acts" (John Searle) characteristically contain referential elements. They refer to something, be it in the physical world of objects, the social world of norms and interaction, or the subjective world of emotion and cognition. Arguably, articulating empirical and theoretical propositions constitutes the nucleus of our linguistic activities. It is what we are supposed to do with social science concepts: developing descriptive and explanatory inferences, making and breaking claims about the social word. In these linguistic performances—our primary speech acts in the social sciences—reference is key. We need our concepts to perform referential roles. We need them to grasp concrete realities in abstract terms. Classical philosophy was centrally concerned with one specific purpose of language: its referential role. We should not be surprised to see that conceptual discussions in contemporary social sciences, too, privilege traditional reference over other linguistic roles.

Meaning

If concepts are means of action, their meaning does not derive from their correspondence to objective realities but from their practical roles in linguistic communication. In Wittgenstein's famous dictum, "The meaning of a word is its use in the language" (1952/1968, sec. 43). Language is a medium of social communication. Its rules of usage and meaning are public, not private. Our shared knowledge about the meaning of a word derives from our shared linguistic practices. As competent language users, we know what others know about the meaning (the conditions of legitimate use) of a concept. As responsible language users, we accept the meaning of a concept (its conditions of legitimate use) when we use it and accept that others can hold us to account for using it. As in other realms of social action, responsibility means that we accept the consequences of our deeds. Take the standard example of a promise. If I promise you x, I understand the meaning of promise making and accept its conditions of validity. Among other things, I understand and accept that x is a future action that benefits you, that I am able to perform it, that my promise obliges me to perform it, and that I actually intend to perform it (Searle, 1969, chap. 3). If I promise, yet violate any of these conditions of validity that constitute the meaning of promises and in consequence fail to carry out x, you can hold me accountable.

When we employ concepts more specifically as means of propositions (the primary form of speech acts in the social sciences), we use them as carriers of general claims about the empirical phenomena they are referring to. When applying them to concrete cases, we subscribe to these claims. We commit ourselves to their truth (applicability). If I call a man a friend, a thief, or a left-wing dictator, I articulate (and thus embrace) certain (contextually understood) claims about my relationship to him, his relationship to alien property, or the form and substance of his exercise of state power. In case of doubt, confusion, or contestation, I must be ready to justify my conceptual choices and accept the consequences. The bundles of claims we commit ourselves to when employing a concept comprises its meaning. Often these claims and commitments remain implicit. Formal definitions serve to make them explicit (Robert Brandom, 2000).

Reification

In political science, we still have to assimilate the insights of pragmatic philosophy. Our discussions of concepts, as far as they take place, still tend to be anchored in the classical distinction between mental creations and real objects. In addition, we tend to reify both sides of the mind–world distinction. We tend to treat both concepts and their referents as if they were things. The result might be described as a kind of double false consciousness. We tend to misrepresent social reality as well as our representations of reality.

(a) The Reification of Reference. Concepts are abstractions, not proper names. They do not serve to designate particular objects but classes of objects. On the referential side, our paradigms of objects are still concrete, material things with observable properties. Very few objects of political research correspond to this model. The realities we study are symbolic. Our concepts are not generalizations from observed properties but abstractions of symbolic realities.

(b) The Reification of Concepts. On the conceptual side, we tend to treat our abstractions, too, as if they were tangible objects, fixed in time and space. Employing the language of factual propositions, we tend to ask what a concept is (and is not) and what its essential attributes are (and are not), as if comprehending the concept required discerning its visible properties.

The Triangle

According to the widely used tripartite conception of concepts developed by Charles Kay Ogden and Ivor Armstrong Richards (the Ogden-Richards triangle), concepts are symbolic entities that consist of three elements: their meaning (connotation or intension), their referents (denotation or extension), and a term or word (their name). This conception of concepts, introduced by Sartori into political science, is still indebted to the notion of concepts as symbolic intermediaries between mind (as location of meaning) and the objective world (as location of reference). It is therefore vulnerable to the reification of both reference and meaning. Still, it serves well to understand the contingent nature of conceptual commitments (the element of choice in the relations between terms, meaning, and reference) and in general offers a set of useful distinctions to analyze the formation and de-formation of concepts (which will guide parts of our subsequent discussions).

Conceptual Disorders

Conceptual discussions in the social sciences often carry therapeutic ambitions. In the pursuit of clarity and precision, they strive to cure scientific language of the multiple disorders that are thought to afflict ordinary language. The tradition of conceptual analysis in political science that was initiated by Sartori and his colleagues in the 1970s and is continued today most prominently by David Collier and John Gerring (2009) partakes of this therapeutic project. Ordinary language is not generally defective, though. It is as clear and precise as speakers need it to be. Still, scientific language differs from ordinary language in some fundamental regards. Among other things, it is written in form and literal in style; it involves a strong commitment to truthful, transparent, and evidence-based argumentation; it aims at generating general knowledge; and it demands the development of a common specialized vocabulary within the academic community. Most of these distinctions are normative, not empirical. They do not give social scientists a mandate to remedy the deficiencies of ordinary language, but they do involve the professional obligation to craft a shared specialized vocabulary that steers clear of major conceptual disorders and malpractices.

Conceptual Opacity

Everyday linguistic communication unfolds on the basis of implicit meaning. Neither do speakers offer formal definitions of the words they use, nor do their interlocutors ask for them—except when their shared understandings turn problematic and when communicative irritations arise, which are instances of incongruence between the concrete application of concepts and their social meaning that we take for granted. You promised to be punctual and are an hour late. Is this your notion of punctuality? You say we are friends, but you left me alone in the face of danger. Is this your idea of friendship?

In the social sciences, we run higher systemic risks of breaking through the thin ice of implicit understandings. Our key concepts are often complex and contested, and we cannot take for granted that others comprehend them in the same manner as we do. Linguistic transparency is, therefore, our first obligation in the social-scientific use of concepts. Karl Marx remarked once that he needed three volumes to explain the concept of capital. We need not go that far in explicating the core concepts we use in our research. Concise formal definitions will often suffice. Yet if we fail to make explicit our central conceptual commitments, our theories and findings cannot contribute to the construction of common knowledge, only to the accumulation of fragmentary statements whose interrelations are uncertain. Conceptual opacity engenders opaque research.

Conceptual Confusion

As the descendents of Noah in ancient Babylonia set out to build a premodern skyscraper ("a tower, whose top may reach unto heaven"), God, alarmed by their capacity of monolingual coordination, intervened to "confound their language, that they may not understand one another's speech" (Genesis 11:7). More than anything else, the concept analytical tradition of Sartori has been concerned about conceptual confusion. According to its disciplinary diagnosis, the builders of comparative political science are afflicted by a similar confusion of tongues as the architects of the Tower of Babel. Lacking discipline and coherence, schools and scholars are speaking past each other with different, mutually incomprehensible vocabularies.

If concepts form triangles of terms, meanings, and referents, conceptual confusion may arise from three sources: confusing relations between terms and meanings (ambiguity), confusing relations between meanings and referents (vagueness), and confusing stipulations of meaning (definitional defects).

Conceptual Ambiguity

To ensure unequivocal associations of words and conceptual commitments "the golden rule is to have one word for each meaning" (Sartori, 1984/2009, p. 113). Confusion may arise if we have several words for one concept (synonyms) or one word with various meanings (homonyms).

Genuine synonyms pose no problem for communication. They enrich our vocabulary and help us avoid tedious repetition. The troubles arise from fuzzy synonyms—neighboring or overlapping terms that are situated in a disorderly semantic field and whose exact relations remain unspecified. Political science is replete with such terms. While language users employ them (loosely) as synonymous, it is unclear whether they actually do carry equivalent conceptual commitments. For instance, *institutions* are often defined as *rules*, and vice versa. Yet we do not know to what extent we can treat the two as interchangeable concepts. In the best of cases, fuzzy synonyms share a recognizable semantic core but differ in their precise connotations (additional shades of meaning). For example, *genocide* and *ethnic cleansing* may refer to the same murderous acts, yet the former maintains analytic distance, while the latter adopts the hygienic discourse of the assassin.

Although words frequently carry multiple meanings, in ordinary language, homonymy does little to disturb our ability to communicate. In everyday talk, context determines meaning. If it fails to do so, we can always ask for clarification. In the social sciences, by contrast, homonymous terms are more corrosive to communication. If scholars attach incongruent meanings to key terms that define their fields of inquiry, they will fail to create common knowledge. Instead of studying one class of phenomena, they will be studying different subjects under the same name. The unity of their field of research will be apparent only, a nominal delusion veiling the substantive fragmentation of their research. For instance, if some hold the goal of "democratic consolidation" to be the prevention of authoritarian regression while others take it to lie in the achievement of democratic deepening, comparative inquiries into the conditions of democratic consolidation will address qualitatively different substantive problems.

Conceptual Vagueness

Concepts are abstractions. They allow us to speak about empirical phenomena in general rather than marvel at them one by one, in puzzlement over their uniqueness (which we could not grasp anyway without a prior notion of generality). In the social sciences, we want concepts to be precise, to circumscribe clearly the realm of phenomena to which they apply. Vague concepts fail to do so. They leave the relation between conceptual claims and empirical objects indeterminate. They do not allow us to decide which phenomena lie inside and which outside their realm of application. Political actors often apply such elastic concepts of unclear denotation as weapons of political struggle. For instance, candidates would condemn “vote buying” by their adversaries without making clear what kind of acts they are referring to. Are they condemning any campaign promise that offers voters material benefits in the future? Squeezed between conflicting expectations, political contenders often benefit from evasiveness. Scholars, by contrast, cannot leave their readers guessing what they are talking about. Leaving its empirical referents in the dark, vague concepts lead our research into obscurity.

Definitional Defects

Our semantic definitions are sources of confusion if their relations to words and things are confusing. They are confusing, too, if their internal structure is defective. Definitional defects may be manifold. Our definitions may be contradictory or tautological. They may be incomprehensible or prone to provoke misunderstandings. They may ignore standard rules of classification by offering classificatory schemes that fail to be exclusive, exhaustive, and one-dimensional. Or they may confound levels of abstraction within taxonomical orders of classes and subclasses. As a matter of fact, there is nothing easier than to be confusing. A moment of distraction, a slip of the tongue, a typographical mistake, and the meaning of what we meant to say dissolves in mist.

Conceptual Instability

In the social sciences, linguistic stability permits continuity in research. The diffusion of new theoretical approaches tends to involve major shifts in our vocabulary. That is what new theories often are: new vocabularies and redescriptions of the world. The waves of conceptual instability they induce tend to reshape collective research agendas. They tend to disrupt established lines of inquiry even when substantive concerns remain essentially the same. Each time we replace our theoretical vocabularies, we tend to reinvent the wheel of empirical research. For instance, the new literature on the political economy of political regime change ignores earlier debates on capitalism and democracy, the new literature on state failure ignores the earlier literature on state building, and so on. Unless we develop and conserve the ability to translate between theoretical languages, destabilizing our core categories can be deeply damaging to the much cherished accumulation of knowledge.

Conceptual Abuse

The tools of language are open to almost limitless forms of abuse. Two symmetrical strategies of reality distortion are of particular interest to political scientists: conceptual stretching (Sartori) and conceptual masking. Conceptual stretching involves the application of (often value-laden) concepts to cases that lack essential characteristics of these concepts (as when a corrupt politician describes himself as an honest man). Conceptual masking involves the description of (morally relevant) cases through neutral concepts that disguise the essential characteristics of these cases (as when a bank robber describes himself as a common customer). The former puts forward semantic overstatements in which concept application betrays fundamental conceptual commitments. The latter puts forward semantic understatements in which concept application denies fundamental factual realities.

These forms of conceptual abuse are surely more frequent and more severe in politics than in political science. Modern authoritarian regimes, masters of linguistic abuse, routinely practice both. Stretching the notion of popular rule beyond recognition, they tend to portray themselves as higher forms of democracy, as when Augusto Pinochet described his form of military dictatorship in Chile as “protected democracy.” The totalitarian regimes of the 20th century were notorious in inventing quasi-neutral technical terms and bureaucratic acronyms to camouflage the unspeakable atrocities they committed against humanity. For instance, the Nazis described their factories of assassination as “concentration camps,” the Soviets under Stalin their officially inexistent colonies of slave labor and death as “Gulag,” “the zone,” or simply and enigmatically “the other side.”

Both malpractices destroy conceptual validity. They sever the link between connotation and denotation, between conceptual commitments and factual applications. In instances of conceptual stretching, speakers claim too much, and realities negate the essence of the concepts they use (often, their moral essence). In instances of conceptual masking, speakers claim too little, and concepts deny the essence of the realities they face (often, their moral essence).

Conceptual Lumping

Names allow us to designate individuals, concepts classes of individuals. More general concepts capture larger classes, more concrete concepts smaller ones. According to the well-known "ladder of abstraction" introduced by Sartori, the number of defining attributes of a concept (connotation or intension) and the number of its referents (denotation or extension) are inversely related. At a high level of abstraction, concepts carry few defining attributes and cover many cases. At low levels of generality, they contain numerous defining features and apply to few cases.

In scholarly research, just as in ordinary language, we have to choose the level of conceptual abstraction that seems appropriate for our purpose. If we talk too abstractly in everyday life, our interlocutors may get irritated: Come down, be concrete, we don't want to hear generalities! If we are overly specific and draw excessively fine distinctions, they may well get impatient, too: Focus on the relevant, stop splitting hairs! In William Ockham's famous formulation, "You need no razor to cut butter."

In crafting social-scientific concepts, we have to seek a pragmatic balance between our ambitions of theoretical generalization and our needs for analytical differentiation. As in everyday interactions, we may err on either side when choosing our levels of conceptual abstraction. If we aim too high and employ excessively general concepts that obliterate "differences that make a difference" (Gregory Bateson), our critics will accuse us of conceptual lumping (Sartori). If we aim too low and choose excessively specific concepts that trace irrelevant distinctions, our critics will reproach us with conceptual splitting.

Overall, to the extent that the social science community is amenable to "linguistic therapy" (Umberto Eco) and avoids developing conceptual pathologies, it strengthens its collective capacity to communicate effectively. Social-scientific language demands more than conceptual health and discipline, though. It also requires theoretical and conceptual creativity, grounded in linguistic competence and empirical knowledge.

Concept Formation

Concept formation is the systematic development and explication of the core claims to which we commit ourselves when applying a concept. It requires us to understand ordinary and specialized uses of the concept, map its location within its semantic field, situate it within empirical realities and analytical frames, understand its structural properties, choose our semantic commitments, and choose the term that best resonates with its meaning.

Reconstruction

If the meaning of a word lies in its use, we need to comprehend the usage of a word if we wish to comprehend the semantic commitments it involves. The first question to ask concerns usage in ordinary language. In English, this question has a straightforward answer: the *Oxford English Dictionary* (OED). Initiated well before the advent of 20th-century pragmatic philosophy of language, the OED is a monument to a pragmatic understanding of language as a medium of social practice. Alien to prescriptive or regulatory pretensions, it meticulously registers "the meaning of everything" (Simon Winchester) by documenting concrete instances of word usage across centuries of linguistic development.

The second question concerns the scientific usage of a word. According to Sartori's seminal "Guidelines for Concept Analysis," the semantic "reconstruction" of a concept starts with a review of the relevant scholarly literature. Unlike the OED, Sartori does not recommend tracing concepts in usage in order to uncover the implicit claims that can be inferred from their practical applications. Rather, he directs scholars to compile lists of explicit definitions (until they get bored by repetition and redundancy), enumerate the attributes included, and bring them "into some meaningful kind of organization" (Sartori, 1984/2009, p. 120). Meaningful organization, indeed, is the central task of concept formation and the most difficult one, as it escapes rules and recipes. To begin with, it requires us to place our concepts in their linguistic contexts.

Semantic Fields

If we wish to understand a concept properly, we must not analyze it in isolation. We need to map the semantic field it inhabits and locate it in the web of relationships that connects it with its conceptual neighbors. Unless we know what the concept shares with proximate concepts and what separates it from them, we cannot grasp its specificity. Neighboring concepts often share large intersecting circles of meaning, yet carry semantic nuances that are relevant for our research purposes. For instance, if we set out to study civil wars, we step into a rich semantic field of overlapping, yet not identical, concepts, such as guerrilla war, revolution, rebellion, ethnic violence, regional violence, organized violence, political violence, state failure, anarchy, political disorder, and political fragmentation. Semantic cartography, surveying and mapping the "systems of terms" (Sartori, 1984/2009, p. 142) that constitute semantic fields, provides a relational and comparative understanding of such clusters of concepts. It helps us grasp better not only the central connotations of interrelated concepts but also their finer shades of meaning that may be decisive in choosing one concept over another.

Empirical Boundaries

In shedding comparative light on the precise meaning (intension) of related concepts, the analysis of semantic fields also helps clarify their precise referents (extension). To the extent that we grasp the differences and similarities in the substantive claims neighboring concepts contain, we grasp the differences and similarities in the empirical phenomena they refer to. For instance, in the semantic field of civil war, some concepts, such as political violence, are situated at high levels of abstraction and include violent actions outside contexts of societal warfare (e.g., terrorism), while others, such as regional violence, refer to more narrow categories of violent conflict. Some concepts, such as ethnic violence, involve claims about the motives of violence, while others do not. Some, such as guerrilla war, emphasize the presence of organized actors, while others, such as political disorder, emphasize the dissolution of central authority. Some concepts, such as state failure, attribute agency to the state, while others, such as civil war, distribute it among societal actors.

All these semantic differences involve empirical differences. They point to different empirical phenomena. Understanding these differences allows us to understand the empirical scope of concepts—their bounded territories. It gives us an idea of what they include and exclude. Depending on what we want to see and what we decide to ignore, it allows us to choose and use the concepts most akin to our analytical interests. Concepts, writes Gary Goertz (2006), are "theories about the fundamental constitutive elements of a phenomenon" (p. 5). Within the semantic field of civil war, we will select our concept of choice depending on the empirical dimensions our theories designate as relevant—the organized nature of civil wars, their political motives, their outcomes, their causal association with state power, or their membership in the general category of political violence. Sometimes, our inherited vocabulary does not trace the precise distinctions we are interested in. Still, our semantic maps allow us to visualize the configurations of empirical boundaries as previous concept users have found them relevant. Though at times incomplete and inconclusive, they open up a first dialogue between concepts and cases, mediated by theory.

Analytic Frames

When scholars reconstruct the scientific usage of their core concepts in relation to proximate concepts (semantic fields), they may come up with long lists of contending or coexisting, divergent or convergent terms, definitions, and applications. In and of themselves, such listings are of little analytic use. The key challenge is to order them in a manner that resonates with theoretical traditions and empirical concerns in relevant fields of research, and this is easier said than done. The standard recommendation sounds simple: First identify underlying analytical dimensions and then show how different concepts and uses of concepts differ along these dimensions. But which are these underlying dimensions? They do not just lie around, self-evident and open to simple inspection. How then can we find them, and how can we construct them?

The answer is perhaps disappointing: Concept formation is not a bureaucratic enterprise but a constructive one (like social science in general). It is not about the mechanical application of rules, be it of logic or language, but about the creative process of abstraction. In our efforts to make sense of divergent definitions and uses of concepts, we have to renounce the comfort of rules. This is the weak point, the structural lacuna, regarding guidelines for concept formation: There is none for this crucial task.

As a first step, it is always helpful to ask about units of analysis. If authors talk about democracy, is it regimes or states they are talking about? If they talk about the rule of law, is it individual decisions or judicial systems they are talking about? Clarity about the kinds of "objects" different usages of a concept strive to grasp often helps us locate them at different levels of abstraction. However, it does not tell us anything about the substantive claims different usages carry. We are not biologists looking at elephants or cows, tangible, objective phenomena whose observable characteristics make them easy prey to the classic logic of classification *per genus et differentiam*. Identifying the general commonalties that determine their genus and then the specific characteristics that determine their species still involves a good deal of complexity and controversy within the biological community and yet in principle requires little "sociological imagination" (C. Wright Mills). But this is not so in the social sciences. If we have determined, for example, that students of "ideology" alternatively refer to structures of thought, language, or behavior, we still do not know what kind of structures they are talking about. We still have to discern (somehow) the analytical dimensions that distinguish different claims about the defining structures of ideologies, such as internal coherence, external differentiation, sophistication, factual accuracy, abstraction, hierarchy, stability, dogmatism, sincerity, and consciousness (Gerring, 1997).

Scholars do not, however, enter the process empty-handed. They are not naive observers of unstructured data. They know the literature, they know the facts and the theories, and they know the language and the paradigmatic cases that define their field of study. The analytical dimensions they introduce (as well as the analytical dimensions their predecessors introduced in the first place) to bring "meaningful order" into conglomerates of contending definitions are anchored in their theoretical and empirical knowledge. It is such theoretical as well as empirical anchorage that makes the meaningful, fruitful, useful organization of semantic fields possible.

Conceptual Structures

Once we have succeeded in bringing analytical order into multiple uses of a concept, we can reconstruct its structural properties. Conceptual structures are configurations of conceptual commitments. Once we have understood the claims authors commit themselves to when using a certain concept, two questions ensue: How strong are their conceptual commitments? And how do those commitments they consider binding (the "essential features" of a concept) relate to each other?

Conceptual Cores

The strength of our commitments to the conceptual claims we articulate varies by degrees. Some claims we hold to be indispensable across contexts. They constitute the core meaning of a concept. Other claims we deem to be secondary and contingent. They form the peripheral and contextual connotations of a concept. According to the classical approach to concept analysis, from Aristotle to Sartori, if we wish to comprehend a concept, we need to identify the former, its semantic core. In the face of multiple uses of a concept, the semantic core is located at the intersecting area of those claims competent concept users declare to be binding ("necessary and sufficient"). Staking out a common ground of binding conceptual commitments ("the core concept") often allows us to distinguish "narrow," "thin," or "minimal" definitions (that limit themselves to the core) from "broad," "thick," or even "maximal" ones (that go beyond, up to envisioning ideal-typical instances of the concept).

Family Resemblances

According to the classical conception of concepts, if different uses of a term do not share common semantic ground, they do not count as instances of the same concept. They appear as instances of multiple concepts. Modern philosophers of language (reinforced by more recent psychological research on typicality effects) have shed doubt on the notion that our ordinary usage of concepts is governed by the strict membership rules of necessary and sufficient conditions. They have contested the notion that concepts carry semantic cores. Rather than committing themselves to a fixed set of indispensable claims, they have suggested that concept users often commit themselves to a more open set of mutually substitutable claims. Wittgenstein (1952/1968, sec. 66–71) introduced the idea of family resemblances to describe concepts that are defined by an ensemble of alternative attributes rather than sharing a core of necessary attributes. In set-theoretic terms, the relevant features of family resemblance or radial concepts do not form intersections but unions. In terms of classical logic, they are not linked by the operator AND (necessity) but by the operator OR (substitutability) (Goertz, 2006, chap. 2).

It seems indeed to be the case that ordinary language users routinely apply empirical concepts, such as fruit and furniture, to concrete objects on the basis of their closeness to typical examples ("prototypes"). These intuitive applications are not based on an invariable set of claims all competent language users subscribe to. However, the use of family resemblance seems less frequent (and less compelling) in the social sciences. When students of politics use the notion of family resemblance, they commonly apply it not to the highest level of abstraction (the definition of general properties of a concept) but to lower levels of abstraction (the definition of constitutive dimensions or the observation of concrete instances of a concept). Concepts described as family resemblances often do seem to share an abstract semantic core (at a high level of generality), even if either their constitutive dimensions or their observational indicators are mutually substitutive (at lower levels of generality).

As a matter of fact, the observation of family resemblances seems to be generally dependent on the prior comprehension (at least vaguely and implicitly) of an abstract semantic core. It is only because we possess a general notion of their common nature that we can discern the elastic observational resemblance of certain classes of cases. Otherwise, we would perceive no more than superficial similarities among disjointed phenomena. Arguably, this is even true for the two paradigmatic concepts of "families" and "games" from which Wittgenstein derived his notion of family resemblances.

Consider families. The sociological literature offers numerous overlapping definitions of families and subtypes of families. For instance, the modern nuclear family typically includes spouses and their dependent children. The notion of family resemblance, however, does not explicate the meaning of the concept of family. It does not respond to the semantic question of conceptual essence but to the phenomenological question of empirical appearance. It does not pretend to identify families and distinguish them from other social groups on the basis of their general characteristics but to identify the members of particular families on the basis of their physical appearance. Family members may look alike, but they are not members of a family (neither in general nor in particular) because they look alike.

Diminished Subtypes

In cases of family resemblance, empirical referents may lack relevant characteristics of a concept and still represent genuine instances of it. In cases of "diminished subtypes" (David Collier & Steven Levitsky, 1997), empirical referents do not fully possess the relevant characteristics of a concept and therefore do not represent genuine instances of it. Classical typologies do not differentiate among members of a category. Objects are either in or out. If they are in, they are equal members of full standing. Diminished subtypes take into account that all members are equal—but some are less equal than others.

Diminished subtypes arise from continuous multidimensional concepts. When all constitutive dimensions of a concept are held to be essential, the full absence of any dimension involves the absence of the phenomenon in question. However, if constitutive dimensions are not dichotomous, but continuous, cases may lie somewhere in between full presence and full absence on any specific dimension. Situated in the "gray zone" (Goertz) of one constitutive dimension, these cases are still recognizable members of a general class of phenomena. Yet they are less than full members. Due to their structural deficiencies, they appear as distant, damaged, distorted representatives. By adding qualifiers (adjectives) to the original concept, we can avoid conceptual stretching (fraudulent claims to full membership in a category) and point to their specific deficiencies. The beauty of diminished subtypes lies in their diagnostic precision.

For example, if democracy requires competitive elections under universal suffrage, the absence of either electoral competitiveness or electoral inclusiveness renders a political regime nondemocratic. Yet governments may impose partial restrictions on either electoral competition or electoral participation that are not severe and systematic enough to involve the absence of either dimension. Such bounded, ambiguous constraints may turn electoral regimes into diminished subtypes of democracy. In the face of bounded restrictions on competition, we might speak of controlled democracies. In the face of bounded suffrage restrictions, we might speak of exclusionary democracies.

Diminished subtypes of multidimensional concepts may arise from the limited presence of one (or some) of their essential dimensions. They may also arise from the full absence of desirable, yet nonessential, attributes. Diminished subtypes of democracy often seem to express structural deficiencies whose benchmarks are not the core of democracy but the ideal of democracy. They refer to democratic regimes that are in full accordance with democratic minimum standards, yet fail to fulfill more stringent expectations to democratic governance. For example, in clientelist democracies, citizens lack the programmatic orientation, and in apathetic democracies, they do not show the participatory enthusiasm we expect from high-quality democracies. In a symmetrical manner, as empirical cases may be underperforming in relation to a particular standard, they may also be overperforming, thus giving rise to "enhanced subtypes."

Contested Concepts

The notion of family resemblances introduces some degree of flexibility and fuzziness into the classical idea of essential attributes. The notion of "essentially contested concepts," formulated by British philosopher Walter Bryce Gallie only a couple of years after the publication of Wittgenstein's *Philosophical Investigations*, involves a potentially more radical objection to the classical idea of conceptual cores. According to Gallie's (1956) seminal paper, all concepts are open to contestation, yet some are "essentially" contested insofar as "their proper use . . . inevitably involves endless disputes about their proper uses" (p. 169). Which is the source of such intrinsic and irresolvable conceptual disputes? There are two plausible answers. One points to the complexities of concept application, the other to inner tensions that lie at the very core of some concepts.

Gallie's (1956) account focuses on concept application. For concepts to become essentially contested, he says, they must be multidimensional—"internally complex"—and normative—"appraisive" (pp. 171–172). Competent speakers recognize and value the various dimensions of the concept. "Any explanation of its worth must therefore include reference to the respective contributions" of each dimension (p. 172). Yet, despite their abstract consensus on the fundamental component parts of the concept, speakers are likely to weight them differently and apply them differently. They are likely to dispute their relative importance as well as their practical implications under changing circumstances. For instance, actors may agree that modern democracy rests on the principles of majority rule and constitutional government but still disagree about their rank order (which is more important) and their operationalization (how they are to be put into practice).

According to a more radical reading, essential contestation affects the very core of concepts, not just their application. It arises when concepts involve irresolvable inner contradictions and when they are founded on genuine dilemmas, paradoxes, or impossible ideals—"conflicting imperatives" (Reinhard Bendix). Consider the ideas of democracy and the rule of law. The etymological root of democracy ("rule of the people") assigns citizens the double, paradoxical role of rulers and ruled. Their powers cannot be maximized simultaneously. In addition, modern democracy arguably embodies fundamental principles (e.g., popular participation and constitutional restraints) that are mutually supportive as well as (at times) mutually subversive. Balancing the trade-offs they impose demands more than establishing simple orders of priorities. The concept of "the rule of law" suggests a mode of impersonal domination where formal rules govern, rather than human beings—the government of laws, not people. Yet the abolishment of human judgment in the exercise of power constitutes an impossible ideal. These inner tensions are constitutive for concepts such as democracy and the rule of law. Since they are irresolvable, they feed an inexhaustible stream of legitimate disputes that are irresolvable as well. In this perspective, appraisive concepts are likely to be essentially contested if they are essentially self-contradictory.

Conceptual Innovation

Language is a social medium of communication and is public and restrained by shared rules of grammar and semantics. At the same time, it is the quintessential medium of personal expression, free and open to variation and innovation as no other societal institution. Ask James Joyce and Ernst Jandl, for example. In the social sciences, our license to innovative uses of language is more limited than, say, in poetry or insult. Still, within constraints, we do enjoy important margins of conceptual innovation.

New Definitions

In the face of contending definitions, scholars cannot rely on the common, accepted usage of a concept. They have to meddle with the rules of the language game by formulating their own conceptual commitments. The particular definitions they develop may relate in various ways to preexisting definitions. They may privilege one usage against others (selective definitions), they may tie new bundles of conceptual claims by accepting some but shedding others (eclectic definitions), they may strive to express underlying commonalities among multiple conceptions (synthetic definitions), or they may embrace new claims that alter the central connotation of the concept (deviant or original definitions). The last move overlaps with the creation of new concepts.

New Concepts

When do we need new concepts? The answer is very simple: when we wish to draw distinctions that we had not drawn before or when we wish to grasp commonalities that we had not grasped before. Sometimes we wish to seize new empirical phenomena, and sometimes we would like to adopt new perspectives on old phenomena. In the study of politics, we are continually confronted with novel realities; continually, we are trying to see together what political actors tend to look at in isolation; and continually, we are developing theories that redefine the relevant boundaries of the political world. Hence, our incessant demand for conceptual creativity.

Like the political realities they try to capture and the political theories they strive to express, new concepts seldom emerge de novo as radical breaks with the past. More often than not, scholars introduce new concepts by modifying old ones. The repertoire of incremental conceptual innovation is broad. Authors may craft new concepts by (a) redefining the substantive meaning of a given concept, (b) importing concepts from other languages or scientific disciplines, (c) remodeling the ladder of abstraction by adding new distinctions or removing established ones, (d) introducing diminished or enhanced subtypes, or (e) changing the property space of a concept (by altering the distance between conceptual poles, shifting their location, or introducing intermediate categories).

New Terms

When we have reviewed the various uses of a concept and its semantic neighbors and when we have understood the configuration of conceptual claims others accept as binding and chosen those we do, we sometimes face the task of "selecting the term that designates the concept" (Sartori, 1984/2009, p. 123)—sometimes but not always. Sometimes we have terminological choices; sometimes we do not.

The names of grand concepts in political science, such as justice, power, and rationality, the terms we use to designate them (and have used to designate them for years, decades, or even centuries), are given and fixed at present. Scholars often disagree about the precise meaning they associate with certain concepts designated by certain terms. They embrace contending "conceptions" of the concepts under dispute. The tight coupling between terms and concepts, however, prevents their semantic disagreements from spilling over into terminological disagreements. Switching the term under discussion would involve switching the concept under discussion. Authors who discuss, for instance, the concepts of justice, power, and rationality either discuss these concepts under these names, or they discuss something else. If they move even to neighboring terms, such as fairness, authority, and intelligence, they move into different (albeit contiguous) conceptual terrains. When terms and concepts are welded together through strong bonds of semantic history and when the former represent the latter without equivalent substitutes, terminological choices precede semantic debates. By choosing a term, we bring the concept it stands for into focus, which then allows us to partake in ongoing disputes about disputed aspects of its meaning.

When concepts are less deeply anchored in history and theory, the sequence can be inverted. We can first determine the substantive claims we are interested in and then settle on appropriate names, either by selecting among available terms or by crafting new ones. If we articulate our conceptual claims in a precise manner, we put ourselves in a position of selecting the precise terms whose connotations correspond most closely to the substance of our concept. Semantic fields that are densely populated with near-synonyms of similar standing (with none of them dominating all the others) offer the most latitude for fine-tuning our terminological choices. For instance, if we study the consolidation of political regimes, we may choose alternative terms that lie in its semantic vicinity yet emphasize diverging substantive concerns. If we wish to stress the duration of regimes over time, we may talk about continuity, endurance, or persistence. If we wish to stress their ability to weather systemic crises, we may talk about resilience, viability, or sustainability. If we wish to stress the process character of consolidation, we may talk about stabilization, institutionalization, or entrenchment.

While refined concepts ask for refined vocabularies, new concepts demand new terms. Whenever the diffusion of new theories (e.g., game theory) or the emergence of new realities (e.g., electoral autocracies) induces waves of conceptual creativity, they are accompanied by waves of terminological innovation. The rise of game theory has brought a whole new vocabulary into political science that includes notions such as backward induction, bounded rationality, perfect equilibrium, incomplete information, cooperative games, focal points, mixed strategies, and so on. The rise of electoral authoritarian regimes since the end of the Cold War has led comparative scholars to propose a broad assortment of labels designed to capture these novel political systems, such as hybrid regimes, semiauthoritarian regimes, inconsistent regimes, multiparty autocracies, competitive authoritarianism, and institutionalized dictatorship. Overall, driven by changing theories as well as changing realities, terminological innovation is a pervasive phenomenon in political science.

Semantic Constraints

Linguistic innovation, in political science as elsewhere, is always anchored in linguistic tradition. While deviating from tradition, it cannot cut itself loose from it. Conceptual innovations in political science, whether they involve new definitions, new concepts, or new terms, are therefore inevitably constrained innovations (whether tightly or loosely). They are constrained by the semantic past of concepts (etymology and conceptual history) as well as by their semantic present (ordinary and specialized usage). They are also constrained by the systemic logic of semantic fields. Modifying individual nodes in the web of interdependent terms that constitutes a semantic field reverberates throughout the entire web. Any pretension to remake our conceptual tools has to recognize this twin linguistic reality that concepts are rooted in their semantic past and present and embedded in their semantic environment. Semantic constraints are not straitjackets. They leave room for critique and creativity, for selective changes and selective continuities—some more, some less. Yet if conceptual innovations depart too sharply from established usage and create arbitrary associations between words and meanings, they will fail to serve as effective tools for communicating novel insights. They will be outright incomprehensible, liable to provoke systematic misunderstandings, or devoid of resonance and thus condemned to oblivion.

Conclusion

"Clear thinking requires clear language," Sartori (1984/2009, p. 102) wrote a quarter of a century ago. To begin with, it requires clear thinking about language. If we learn to incorporate conceptual self-awareness into our canon of methodological sophistication, we will do better theory and better research. We may not reach the impossible ideal of a fully transparent, clear, and precise technical language, and as a scientific community, we may be too diverse and competitive to build another tower of Babel. Yet we will put the edifice of political science on more solid foundations.

Andreas Schedler
Centro de Investigación y Docencia Económicas (CIDE)
Mexico City, Mexico

See also Comparative Methods; Cross-National Surveys, Discourse Analysis; Hermeneutics; Measurement; Methodology; Political Communication; Political Philosophy; Positivism; Qualitative Comparative Analysis; Thick Description

Further Readings

Austin, J. L. (1962). *How to do things with words*. Oxford, UK: Oxford University Press.

Bevir, M., & Kedar, A. (2008). Concept formation in political science: An anti-naturalist critique of qualitative methodology. *Perspectives on Politics, 6*(3), 503–517.

Brandom, R. B. (2000). *Articulating reasons: An introduction to inferentialism*. Cambridge, MA: Harvard University Press.

Collier, D., & Gerring, J. (Eds.). (2009). *Concepts and method in social science: The tradition of Giovanni Sartori*. New York: Routledge.

Collier, D., & Levitsky, S. (1997). Democracy with adjectives: Conceptual innovation in comparative research. *World Politics, 49*(3), 430–451.

Gallie, W. B. (1956). Essentially contested concepts. *Proceedings of the Aristotelian Society, 56,* 167–198.

Gerring, J. (1997). Ideology: A definitional analysis. *Political Research Quarterly, 50*(4), 957–994.

Goertz, G. (2006). *Social science concepts: A user's guide*. Princeton, NJ: Princeton University Press.

Murphy, G. L. (2004). *The big book of concepts*. Cambridge: MIT Press.

Sartori, G. (2009). Guidelines for concept analysis. In D. Collier & J. Gerring (Eds.), *Concepts and method in social science* (pp. 97–150). New York: Routledge. (Original work published 1984)

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge, UK: Cambridge University Press.

Wittgenstein, L. (1968). *Philosophical investigations* (3rd ed.; G. E. M. Anscombe, Trans.). New York: Macmillan. (Original work published 1952)

Conditionality

Conditionality in political science refers to the use of the conditions an actor attaches to awarding benefits to—or to not imposing costs on—another actor in order to influence his behavior. Benefits are granted only if the target actor has fulfilled certain requirements (*ex ante* conditionality). Such external rewards can also be suspended or withdrawn if the recipient reneges on his commitment (*ex post* conditionality). Rather than withdrawing a promised or granted reward (positive conditionality), the violation of predetermined conditions can also be punished (negative conditionality). Conditionality targets either performance and outcomes (policy conditionality) or the governance and administrative systems (political/democratic conditionality).

The term refers more narrowly to the Bretton Woods institutions seeking to impose the Washington Consensus around the world or, more broadly, to all explicit and implicit requirements for lending, including covenants for project-based aid, environmental safeguards, and performance-based aid allocation. This entry follows the broader understanding of conditionality as a political instrument used by national governments and international organizations to advance their foreign policy goals. First, the concept is linked back to the instrumental logic of social action in order to specify the causal mechanisms through which conditionality is to influence the behavior of third countries (i.e., those who are not members of the European Union [EU]). Second, the emergence of conditionality as a key instrument of development cooperation and foreign policy, more broadly speaking, is traced. Finally, critical evaluations regarding the effectiveness and legitimacy of conditionality are discussed to outline how these have led to readjustments of the instrument.

Conditionality and the Manipulation of Utility Calculations

Conditionality seeks to influence the behavior of actors through the manipulation of utility calculations, by providing negative and positive incentives. Donor states or organizations seek to promote structural adjustments in third countries by promising or granting additional benefits, such as financial and technical assistance, a loan, debt relief, or conditional membership in an organization. Or they incur costs through economic and diplomatic sanctions. "Reinforcement through reward" (Frank Schimmelfennig & Ulrich Sedelmeier, 2006) or "correction through punishment" differs from political instruments based on a normative logic of social action, such as political dialogue, which seek to change the behavior of actors through persuasion and learning (Jeffrey T. Checkel, 2005). Conditionality and political dialogue both aim at influencing the choice of actors, whether they are informed by cost–benefit calculations or guided by normative concerns about socially accepted behavior. They thus contrast with coercion and assistance. While the former does not leave actors any choice, the latter provides unconditional financial and technical aid to enable actors to make choices.

Evolution of Conditionality as an Instrument of External Governance

Conditionality emerged in the field of development cooperation. In the 1980s, the International Monetary Fund and the World Bank increasingly linked loans, debt relief, and financial aid to certain conditions, known as structural adjustments, which included aid effectiveness (e.g. anticorruption measures), financial austerity, or the privatization of key public services. While international donors have used conditional aid to promote selective reforms deemed necessary for good governance, the EU turned conditionality into the key instrument in its attempt to support the transformation in postcommunist countries after the end of the Cold War (Heather Grabbe, 2006). The Maastricht Treaty had made any contractual relations of the EU with third countries conditional on the respect for human rights, the rule of law, and democracy (see Tanya Börzel & Thomas Risse, 2009). States applying for membership in the EU have to fulfill the so-called Copenhagen Criteria (democratic conditionality). Once the EU agreed to open accession negotiations, accession conditionality gave the European Commission a powerful tool to push candidate countries toward downloading the comprehensive *acquis communautaire* and introducing institutional reforms (see Schimmelfennig & Sedelmeier, 2006).

Conditionality is also applied in the EU's attempt to transform the domestic structures of its neighbors, which do not have a membership perspective (yet). The European Neighbourhood Policy offers the Mediterranean countries and the Western Newly Independent States financial and technical assistance, progressive integration into the internal market and its regulatory structures, and closer diplomatic relations (e.g., association) in return for economic and political reforms. Conversely, it may (threaten to) suspend bilateral agreements, withhold assistance, and impose political sanctions (e.g., visa bans). Donor–recipient relationships in the European Neighbourhood Policy are less tight than in EU enlargement policy, but association agreements still create closer ties than most international instruments.

Evaluation and Adjustments of the Instrument

The efficiency of conditionality has become a key concern for analysts and practitioners. The approach as developed in the 1980s has increasingly been called into question. There is wide consensus over the limited effectiveness of both policy and political conditionality in development cooperation. Donor leverage cannot substitute for domestic political will, leading to the call for stronger government ownership and *ex post* evaluation of outcomes. Research on EU conditionality points to similar limitations regarding the impact of conditionality. While target states have formally adopted a massive amount of EU legislation, this is often not properly applied and enforced and, thus, has not changed the behavior of actors (Gerda Falkner, Oliver Treib, & Elisabeth Holzleithner, 2008). Weak norm internalization has given rise to concerns about shallow Europeanization. As in development cooperation, the limited administrative capacities of the candidate countries have curbed the domestic impact of EU accession, which accounts for its differential outcome (Schimmelfennig & Sedelmeier, 2006).

Overall, the effectiveness of conditionality seems to depend on a number of specific scope conditions. The promised rewards have to be credible and big enough to pay off the compliance costs; the same applies to the threat of sanctions. Moreover, there needs to be a coalition of actors within the target country that is willing to use external incentives to push for compliance domestically (Jan Pronk, 2001). These conditions are seldom met, even among the current EU candidate countries, including the Western Balkans and Turkey.

In reaction to the limited effectiveness of conditionality, issues of recipient ownership, greater selectivity based on country performance, stronger results orientation, and thus a shift to *ex post* conditionality have gained relevance. While the World Bank and the International Monetary Fund aim to streamline conditionality and rely on greater selectivity and new approaches, such as programmatic lending, the European Commission attempts to reconcile a predictable budget support for low-income countries with a stronger performance orientation that bases disbursements on outcomes. Other critical observers call for the abandonment of conditionality altogether and argue instead for broad political dialogue and the allocation of funds on the basis of need.

The legitimacy of conditionality is equally contested. On the one hand, the output legitimacy of external interference has come under fire because structural adjustments have failed to produce the intended outcomes or have even been blamed for causing negative effects. On the other hand, the input legitimacy of conditionality is challenged by international norms of state sovereignty and self-determination. This legitimacy problem is exacerbated if externally promoted norms and standards are not equally promoted inside the donor organization or state (double standards). At the same time, conditionality has a legitimating function inside the donor states to justify payments to third countries, especially under high uncertainty regarding the effect on external actors. There seems to be a trade-off between legitimacy concerns in recipient and donor countries that creates a serious normative dilemma and adds to the controversy surrounding the use of conditionality both among policymakers and researchers.

Tanja A. Börzel
Freie Universität Berlin
Berlin, Germany

Eva G. Heidbreder
Hertie School of Governance
Berlin, Germany

See also Democratization; Developing World and International Relations; Development, Political; European Integration

Further Readings

Börzel, T. A., & Risse, T. (2009). Venus approaching Mars? The European Union's approaches to democracy promotion in comparative perspective. In A. Magen, M. McFaul, & T. Risse (Eds.), *Democracy promotion in the US and the EU compared*. Houndmills, UK: Palgrave Macmillan.

Checkel, J. T. (Ed.). (2005). International institutions and socialization in Europe [Special issue]. *International Organization, 59*(4).

Falkner, G., Treib, O., & Holzleithner, E. (2008). *Compliance in the European Union: Living rights or dead letters?* Aldershot, UK: Ashgate.

Grabbe, H. (2006). *The EU's transformative power: Europeanization through conditionality in Central and Eastern Europe*. Houndsmills, UK: Palgrave Macmillan.

Pronk, J. P. (2001). Aid as a catalyst. *Development and Change, 32*(4), 611–629.

Schimmelfennig, F., & Sedelmeier, U. (2006). Candidate countries and conditionality. In P. Graziano & M. P. Vink (Eds.), *Europeanization: New research agendas*. Houndsmill, UK: Palgrave Macmillan.

Conditions, Necessary and Sufficient

Social science, as all science, is a continuous quest for an explanation and understanding of the world around us. This search is carried out through Benjamin Most and Harvey Starr's "research triad" of theory, logic, and research design, all of which are central to both hypotheses and results. The key element of explanation and understanding is causation. Social scientists are concerned with causation as applied both to individual events or cases and to classes or groups of events. The causal relationship may take many forms. Two of the most prominent, important, and commonly used forms of the causal relationship involve necessary and/or sufficient relationships. As emphasized in the work of Most and Starr among others, analyses must be concerned with the form of the relationship. David Hume's classic definition of cause involves the "constant conjunction" of an object followed by another, where all objects similar to the first are followed by objects similar to the second. Causation is seen as including three elements, the first being the existence of correlation between two factors or variables. That is, changes or attributes in one factor are associated with changes or attributes in another. Correlation is one possible form of any specific relationship. Correlation itself may take several forms, such as linear or curvilinear. The second element of cause involves the temporal dimension—the proposed causal factor must take place before the "effect" or the phenomenon to be explained. Cause must precede effect. The third element is the most rigorous and difficult—the elimination of other explanatory factors beyond those proposed (the central idea of "control"). Below, the various implications of these relationships and some current applications in political science are discussed.

Necessity and sufficiency are themselves different forms of the causal relationship. They represent different forms of constant conjunction. Scholars have argued that research designs (the nature of the theoretical logic and research hypotheses being employed) built around these different forms of causation will be affected by the nature and types of the cases selected, the controls employed for dealing with possible other (explanatory) factors, and the methods used for evaluating the theory and proposed research hypotheses. All three elements of the research triad—theory, logic, and research design—are thus affected by the form of the relationship, especially looking at necessary relationships in distinction from sufficient relationships or even simple correlational relationships. The differences between how necessary and sufficient conditions are treated are additionally important in regard to the research designs needed to investigate inference in the small-*N* studies that characterize much of comparative politics (and qualitative analyses more generally). The logics of necessary and sufficient conditions need to be investigated in order to cross the boundaries between political science subfields as well as the quantitative–qualitative divide.

Definitions and Logics

The most basic definition of sufficiency states, "if *X* then *Y*." That is, *X* always leads to *Y*, but *Y* is not always preceded by *X*. For example, proponents of the theory of democratic peace argue that democracies have peaceful relations with one another—that is, a pair of democracies (X) will lead to peace (Y), which is defined as the absence of war. However, peace may result from many other factors (e.g., power preponderance, lack of contact or opportunity), so that it may occur without the presence of a democratic dyad. The presence of democratic dyads predicts the presence of peace: *X* "yes" and *Y* "yes," as in Cell A of Figure 1. A much less important prediction (or expectation) is that the absence of a democratic dyad is followed by an absence of peace: the "no/no" Cell D of Figure 1. Cell B in Figure 1, where a democratic dyad is not present but peace is, is not relevant for sufficiency, because *Y* may occur without *X* (i.e., when *X* is a sufficient but not necessary condition for *Y*). The key relationship is in Cell B, when *X* does occur but *Y* does not. This cell will be empty when *X* is a sufficient condition for *Y*. Thus, a research design that looks for sufficient relationships must look at the full range of *Y*, the dependent variable—when *Y* occurs and when it does not. Herein rests the strong admonition by scholars that the researcher cannot select on the dependent variable (which would toss out the ability to look at the key cell that permits the investigation of sufficiency).

	$X = 1$	$X = 0$
$Y = 1$	(A) Predicted	(B) Irrelevant
$Y = 0$	(C) KEY—should be empty	(D) Predicted

Figure 1 *X* Is Sufficient for *Y* (If *X* Then *Y*)

	$X = 1$	$X = 0$
$Y = 1$	(A) Predicted	(B) KEY—should be empty
$Y = 0$	(C) Irrelevant	(D) Predicted

Figure 2 *X* Is Necessary for *Y* (Only if *X* Then *Y*)

In many ways, necessity produces reverse expectations. Necessity means the following: only if *X*, then *Y*. That is, *Y* is always preceded by *X*, but *X* does not always lead to *Y*. Bruce Bueno de Mesquita's theory of positive expected utility (EU) notes that positive EU is a necessary condition for war—that war (*Y*) is always preceded by positive EU on the part of decision makers going to war (*X*). In essence, necessity asks us to look for preconditions or prerequisites of some phenomenon *Y*. It should be noted that the authors in Gary Goertz and Harvey Starr's edited volume define necessity in various ways and using different logics: set theory, calculus, a probabilistic logic, fuzzy logic/sets, and Aristotelian two-valued logic. Returning to the current example, simply having positive EU does not mean that there will be war, for a variety of domestic and external factors, again including lack of opportunity. As seen in Figure 2, Cell A, the presence of the causal factor (positive EU), predicts or is followed by the presence of war, or the dependent variable. A much less important prediction (or expectation) is that the absence of the independent variable or causal factor *X* is related to the expectation or prediction that the effect, or dependent variable, *Y* is also absent (Cell D). Cell C presents the condition where *X* is present, but there is no effect on *Y*. This is not relevant to necessity because *X* does not always lead to *Y* (indeed, this has been a major confusion in critiques of the early EU models of Bueno de Mesquita and colleagues). For necessity, the key relationship is when the causal factor *X* does not occur but the effect *Y* still appears—Cell B. If *X* is a necessary condition for *Y*, this cell or set will be empty. Note that this means that the whole range of the independent variable *X* must be included. One must see cases where *X* does not occur. Necessity then can be tested using only the occurrence of *Y*, the dependent variable. With questions of necessity, therefore, the researcher may select on the dependent variable but cannot select only on the occurrence of *X*, the independent variable. This distinction between sufficiency and necessity suggests that case selection should be guided by the outcome that the researcher is interested in and how the logic of the proposed causal mechanism identifies that outcome (as argued by Goertz and colleagues). As noted above, one evident way to do this is with set theory. Thus, a strong set-theoretic basis to the logic of necessary relationships is found.

It should be clear that the logics of necessity differ from those of sufficiency. Much of the analyses performed on relationships of sufficiency use continuous variables, looking to find linear relationships on normally distributed values and to maximize the explanation of variance in the dependent variable (these have been called frequentist statistical methodologies). Necessity more often looks at the presence or absence of an explanatory factor *X*. That is, a necessary analysis looks to see if values on a measure have crossed a threshold, reached a "turning point," or have come to a "critical juncture" (see also the "powder keg" explanations discussed by Gary Goertz and Jack Levy). Goertz thus argues that necessary conditions and the statistical models related to them set out different sorts of hypotheses from standard quantitative analyses of sufficient relationships. Rather than tracking how changes in any *X* are related to changes in some *Y*, most necessary analyses focus on whether *X* makes *Y* possible or more likely. Goertz contrasts a "necessary condition hypothesis," which might propose that some minimum level of variable *X* is necessary for an outcome variable *Y*, with a "correlational hypothesis," which would propose that the greater the value of some *X* then the greater the value of some *Y*. Another example of analytical differences has been identified by Charles Ragin, among others. He has argued that while standard statistical techniques are often applied to deal with (i.e., remove) heteroskedasticity, this is a central condition and characteristic of necessary relationships and thus needs to be part of the analysis and not removed.

Necessary and sufficient hypotheses, then, often work within different causal chains. Correlational hypotheses tend to focus on classes of events, while necessary-condition hypotheses tend to focus on causation within a single case (but which may also be aggregated probabilistically across such cases). The latter is often associated with process tracing within cases to identify the causal mechanisms at work and is thus of importance to comparative politics, qualitative approaches, and comparative-historical methodology. Goertz has demonstrated the range, ubiquity, and importance of necessary-condition hypotheses found throughout the political science literature. They often take the form of counterfactuals, especially in case studies focusing on explaining individual events. The necessity relationship claims that *only if X* can there be *Y*. The counterfactual claim holds that *without X*, if *X* had not happened, *Y* would not have occurred. Goertz and Levy contrast a necessary-conditions counterfactual approach (looking for necessary conditions within individual cases) with a nomological covering-law approach (based on the constant conjunction of covering laws stated and tested with large-*n* statistical/probabilistic methods). More than 2 decades ago, Most and Starr demonstrated that researchers will produce meaningless findings when creating research designs that can test only for necessary conditions while stating their hypotheses in terms of sufficiency (and vice versa). That is, the cases selected and the methods used must match the questions asked. The method of selecting cases will logically preclude some questions and some designs.

Scholars across political science subfields as well as different research traditions have come more and more to seek designs that uncover complex causality, recognizing that there may be multiple causal paths to the same outcome. This is often called *equifinality*. Contingency plays a large part in such causal complexity—that some theory or hypothesis will hold only under certain conditions or that a pair of contending theories may both be true under different conditions or contingencies (or what Most and Starr call "nice laws"). In many cases, contingency is expressed as a necessary condition, without which some effect, outcome, or other dependent condition could not occur. This is why necessity is a key component of a philosophical view of causality becoming more prominent in the literature—the INUS view of causation. An INUS explanation is an "Insufficient but Nonredundant [i.e., necessary] part of an Unnecessary but Sufficient condition," for which James Mahoney and Gary Goertz (2006, p. 232) provide a succinct explanation:

> An INUS cause is neither individually necessary nor individually sufficient for an outcome. Instead it is one cause within a combination of causes that are jointly sufficient for an outcome. Thus, with this approach, scholars seek to identify combinations of variable values that are sufficient for outcomes of interest. The approach assumes that distinct combinations may each be sufficient, such that there are multiple paths to the same outcome.

Research findings with INUS causes can often be formally expressed through Boolean equations such as Y = (A AND B AND C) OR (C AND D AND E).

The logic of necessity and sufficiency, and multiple causal paths, may thus be clearly captured by Boolean logic and methods (most fully introduced in the work of Charles Ragin). The conjunctive Boolean AND denotes necessary conditions. Thus, in the above equation, A, B, and C are all necessary conditions in combination, to be sufficient for Y. But note, so are C, D, and E in combination. The Boolean disjunctive OR indicates that either of these two combinations is sufficient for *Y*—multiple paths to the same outcome. The reader should note that the philosophy of science dealing with concepts holds a similar position—that a good definition of a concept is one with necessary conditions that are jointly sufficient.

Conclusion

This brief overview has attempted to present the most recent reflections and findings on necessity and sufficiency, setting out the basic phenomena and issues that are currently being discussed in the political science literature. The reader who investigates the works cited and suggested below and follows the footnotes and references in those materials will obtain the most comprehensive picture of current thinking on necessity and sufficiency. These works reinforce the observation of Most and Starr that there is a need for a better understanding of logical linkages that serve as alternatives to the standard statistical techniques commonly employed in large-*N* statistical analyses with an almost exclusive focus on sufficient relationships.

These newer works provide the theory, logic, and findings that support Mahoney's refutation of a set of "erroneous beliefs" about necessity and sufficiency. Based on the present discussion and the works cited, one can, in fact, say that

1. necessary and sufficient causes do exist and are found in interesting hypotheses;
2. necessary and sufficient causes are not trivial, tautological, or irrelevant (see Goertz's work for an in-depth conceptualization of the trivialness, importance, and relevance of necessity and sufficiency and ways to measure them);

3. necessary and sufficient causes are neither deterministic nor inconsistent with probabilistic analysis, and they can be measured continuously (see especially Ragin); and
4. subsequently, methods do exist to test hypotheses of necessity or sufficiency, including tests for significance/trivialness.

For Points 3 and 4, see especially work by Mahoney as well as Goertz.

Finally, and perhaps most important, newer works on necessity and sufficiency point out ways in which research can be designed to capture causality (in its various forms) across small-*N* and large-*N* analyses and cross boundaries between so-called qualitative and quantitative analysis. In so doing, this work also enhances boundary crossing between international relations and comparative politics at a time when greater and greater attention is being directed toward two-level analysis, taking into account the impact of internal/domestic factors and context along with external/international factors and context.

Harvey Starr
University of South Carolina
Columbia, South Carolina, United States

See also Case Studies; Causality; Comparative Methods; Hypothesis Testing; Quantitative Methods, Basic Assumptions; Quantitative Versus Qualitative Methods

Further Readings

Goertz, G. (2006). Assessing the trivialness, relevance, and relative importance of necessary or sufficient conditions in social science. *Studies in Comparative International Development, 41,* 88–109.

Goertz, G., & Levy, J. S. (Eds.). (2007). *Explaining war and peace: Case studies and necessary condition counterfactuals.* New York: Routledge.

Goertz, G., & Starr, H. (Eds.). (2003). *Necessary conditions: Theory, methodology and applications.* Lanham, MD: Rowman & Littlefield.

Mahoney, J., & Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis, 14,* 227–249.

Most, B. A., & Starr, H. (1989). *Inquiry, logic and international politics.* Columbia: University of South Carolina Press.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy-sets and beyond.* Chicago: University of Chicago Press.

Configurational Comparative Methods

Configurational comparative methods (CCMs) designate both a methodological approach and a set of specific techniques enabling systematic cross-case comparison while also taking into account within-case complexity. They have been developed, initially, for small- and intermediate-*N* research designs in the social sciences (Benoît Rihoux & Charles Ragin, 2009). The whole CCMs enterprise was initiated by Ragin (1987), who outlined a "synthetic strategy" that could bridge case-oriented and variable-oriented approaches to social phenomena. He also developed a first technique, QCA (qualitative comparative analysis—now labeled *crisp-set QCA* [csQCA]), and applied it to macro-level cases such as countries. Below, the basic assumptions, specific techniques, and applications, and recent developments are discussed.

Core Assumptions and Goals

The whole ambition behind CCMs is, on the one hand, to make qualitative case analysis more systematic and to offer case-oriented researchers some tools that enable them to systematically compare thick, complex cases. On the other hand, it strives to offer an alternative way to envisage causal arguments made in mainstream quantitative (read statistical) social-scientific work. The overarching goal of CCMs is to unravel causal complexity by applying set-theoretic methods to cross-case evidence. In more concrete terms, the different QCA techniques developed within CCMs enable one to identify core combinations of conditions (input variables), which explain the variation of a given outcome (output variable) of interest. Therefore, QCA techniques are geared toward the identification of so-called specific connections between conditions and outcomes. In contrast to most statistical techniques, they are not geared toward the establishment of general, tendential, or correlational connections between each independent variable, on the one hand, and the dependent variable, on the other (Ragin, 2008; Rihoux, 2008).

Causality, or linkages between conditions and the outcome, more generally, are assumed to be multiple and conjunctural. This implies the following assumptions: (a) Most often, it is a combination of conditions (rather than a single condition) that generates the outcome; (b) several different combinations of conditions may produce the same outcome; and (c) there is no fixed effect of a given condition on the outcome—depending on how it is combined with other conditions, different values of this condition can produce the outcome. The concrete goal of CCMs, from this perspective, is to identify those different causal paths leading to some outcome of interest, each path being relevant in its own way, regardless of the number of cases it covers. Thus, this concept of causality does not take on board most of the core assumptions underlying mainstream statistical analysis, such as permanent causality, uniformity of causal effects, unit homogeneity, additivity, linearity, and causal symmetry (Dirk Berg-Schlosser, Gisèle De Meur, Benoît Rihoux, & Charles Ragin, 2009).

The logical foundations of CCMs can be found in John Stuart Mill's "canons," in particular the method of agreement, the method of difference, and the joint method of agreement and difference, which are all logical ways to systematically contrast and match cases, so as to establish common causal relationships. In practical terms, in CCMs, this is translated into tests for necessity and sufficiency.

Techniques

The overall rationale behind the QCA techniques is that, through some step-by-step logical operations (based on Boolean algebra or set-theoretical logic), one is able to reduce complex data tables to shorter combinations of conditions explaining the outcome of interest. Therefore, at the heart of CCMs lies the quest for parsimony—but not in such a way that within-case complexity would be sacrificed. So far, three specific techniques have been developed under the umbrella term of *QCA*: crisp-set QCA (csQCA), the initial technique that uses Boolean, that is, dichotomous, sets; multivalue QCA (mvQCA), which allows the use of multiple-category conditions; and fuzzy-set QCA (fsQCA), which uses finer grained fuzzy-set membership scores.

These techniques can be exploited in at least five different ways. The first two uses are more descriptive or classificatory. One can use QCA techniques, first, simply to summarize data through the production of a truth table (or table of configurations), which can be examined to discover, for instance, how cases cluster together, or to elaborate typologies. Second, QCA techniques can be used as a tool to assess the coherence of the data—in particular through the detection of so-called contradictory configurations—which lead the researcher to reflect on the cases and the theory.

The three other, more full-fledged uses of QCA techniques imply some sort of testing of a causal argument: checking hypotheses or existing theories, quickly testing some conjectures made by the researcher, and developing new theoretical arguments. Thus, QCA techniques can be used for both theory testing and theory building.

In a full QCA application, whatever the specific technique, the core steps are basically the same (Benoit Rihoux & Bojana Lobe, 2009). In a first phase, upstream of QCA proper, a comparative research design is chosen, cases are selected in a purposeful way, sufficient intimacy is gained with each one of the cases, the outcome of interest is defined, the conditions are selected (model specification), and the cases and the model are visualized or synthesized in various ways. These first operations already result in some level of reduction of complexity by transforming thick, complex case narratives to analytical (i.e., variable-based) data.

The second core phase corresponds to the more formal, computer-run part of QCA. There are several steps to be performed: operationalizing the outcome and condition variables (dichotomization or some finer grained calibration), producing the truth table, solving the contradictory configurations so as to obtain a contradiction-free truth table, minimizing the truth table (with and without the inclusion of logical remainders, i.e., nonobserved cases), solving contradictory simplifying assumptions produced by the use of logical remainders, and finally obtaining minimal formulae that contain those core, most parsimonious combinations of conditions leading to the outcome of interest.

The third phase, downstream, consists of different ways of moving back to more (case) complexity: identifying more crucial conditions, interpreting the minimal formulae in a case-by-case manner, interpreting cross-case patterns, performing limited historical or modest generalizations, and finally, possibly, going through the next cycle of

QCA, for example, through the examination of some more, rather proximate cases.

All these steps should be performed in a reflexive, iterative way, and with frequent loops back to theoretical and/or case-based knowledge. As QCA techniques have been consolidating and as applications have been becoming more diverse, a body of good practices has also been consolidated.

Applications

The applications of CCMs have become increasingly diverse. More than two thirds of them are found in various subfields of political science—in particular, comparative politics and policy studies. The breadth of applications has expanded in several respects. In terms of number of cases, though QCA techniques were initially geared toward small- and intermediate-*N* designs, larger-*N* applications have also proved successful. In terms of the nature of cases envisaged, most applications so far envisage macrolevel (e.g., countries, political systems, and policy processes) or mesolevel (e.g., party organizations, social movements, and collective actors) cases. However, some researchers are beginning to apply QCA on microlevel cases as well. Another type of broadening is related to the fact that some users who are more application oriented are also becoming interested in such methods, for example, in the field of policy evaluation and monitoring, beyond the academic sphere. With regard to the number of conditions included in the analysis, there is also broad variation, though a good practice dictates that QCA models be kept relatively short due to the limited-diversity issue—that is, too long a model will not be in the interests of parsimony.

Gradually, applications have become technically more sophisticated. Two main computer programs have been developed and are now used for most applications: TOSMANA (Tools for Small-*N* Analysis; for csQCA and mvQCA) and FSQCA (fuzzy-set qualitative comparative analysis for csQCA and fsQCA). Some other interfaces are also being gradually exploited, such as R and STATA.

Debates and Further Developments

As CCMs are still rather recent and as they are being perceived by some as a challenge to some other methodologies, they are subject to a lot of debate. The core critiques concentrate in particular on the following points: the dichotomization of the data, the use of nonobserved cases, case sensitivity, the difficulty in selecting conditions, the inability of CCMs to actually establish causal linkages and broad generalizations (inference), and the nonexplicit inclusion of the time dimension. Naturally, as with all methods, there are limitations and weaknesses with CCMs, but many strategies have been developed to overcome most of these drawbacks. One should also bear in mind the specific goals and assumptions of CCMs. For instance, contending that CCMs are weaker than statistical techniques because they do not produce inference (generalization from a sample to a whole population) is a misplaced criticism, because this is not the goal of CCMs—they follow a more case- and diversity-oriented logic.

Many developments and improvements of CCMs are under way—the launching of mvQCA and fsQCA is an example; these programs were developed as a response to some of the limitations of the dichotomous csQCA. Among the many ongoing refinements, one can mention the following: the improvement of computer software, the inclusion of the time dimension, the distinction between distant and proximate conditions (two-step models; see Carsten Schneider & Claudius Wagemann, 2006), the Most Similar Different Outcome/Most Different Same Outcome (MSDO/MDSO) procedure in the process of model building (Berg-Schlosser & De Meur, 2009), more elaborate strategies to tackle logical remainders and contradictory simplifying assumptions, better visualization of the cases and the minimal formulae, and so on.

In another line of improvements, CCMs have been increasingly combined or sequenced with qualitative and/or quantitative methods. On the smaller-*N* side of the spectrum, the connection with qualitative, case-oriented work is self-evident, as in-depth case knowledge goes along with the QCA protocol—the graphic display of synthetic case descriptions is one way to bridge more explicitly the thick case evidence and the formal QCA procedure (Rihoux & Lobe, 2009). In intermediate- and larger-*N* designs, several works have fruitfully confronted or mutually enriched some QCA techniques with statistical techniques, such as discriminant analysis, factor analysis, and various types of regression-based analyses.

Many more developments can be expected within the next few years as the community of users of CCMs is enlarging and also becoming more diverse in disciplinary terms—in fact, some of the cutting-edge work is now also being done in fields such as management, organizational studies, and communication studies. The speed of further improvements of CCMs will largely depend on the ability of this growing community to share experiences as well as innovations and to make this a cumulative enterprise.

Benoît Rihoux
Université Catholique de Louvain
Louvain-la-Neuve, Belgium

See also Case Studies; Causality; Comparative Methods; Complexity; Conditions, Necessary, and Sufficient; Fuzzy-Set Analysis; Qualitative Comparative Analysis; Quantitative Versus Qualitative Methods; Quantitative Methods, Basic Assumptions

Further Readings

Berg-Schlosser, D., & De Meur, G. (2009). Comparative research design: Case and variable selection. In B. Rihoux & C. C. Ragin (Eds.), *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (pp. 19–32). Thousand Oaks, CA: Sage.

Berg-Schlosser, D., De Meur, G., Rihoux, B., & Ragin, C. C. (2009). Qualitative comparative analysis (QCA) as an approach. In B. Rihoux & C. C. Ragin (Eds.), *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (pp. 1–18). Thousand Oaks, CA: Sage.

COMparative methods for the advancement of systematic cross-case analysis and small-n studies. Retrieved November 10, 2010, from http://www.compasss.org

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago: University of Chicago Press.

Rihoux, B. (2008). Case-oriented configurational research using QCA (qualitative comparative analysis). In J. Box-Steffensmeier, H. Brady, & D. Collier (Eds.), *Oxford handbook of political science: Methodology*. Oxford, UK: Oxford University Press.

Rihoux, B., & Lobe, B. (2009). The case for QCA: Adding leverage for thick cross-case comparison. In D. Byrne & C. C. Ragin (Eds.), *The SAGE handbook of case-based methods* (pp. 222–243). London: Sage.

Rihoux, B., & Ragin, C. C. (2009). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. Thousand Oaks, CA: Sage.

Schneider, C. Q., & Wagemann, C. (2006). Reducing complexity in qualitative comparative analysis (QCA): Remote and proximate factors and the consolidation of democracy. *European Journal of Political Research, 45*(5), 751–786.

Conflict Resolution

Conflict resolution is a term that we all understand until we try to define it. Synonyms abound: such as *conflict settlement*, *conflict termination*, *conflict management*, and *conflict transformation*. Moreover, conflict resolution is both a situation and a process. A conflict is resolved when all the parties to a dispute agree that it is over for good, in full knowledge of the situation, and without any form of coercion, whether personal, manifest or structural, since they acknowledge that their respective interests and values have been satisfied. The concerned parties are those who cannot be seen off or defeated and without whom there can be no resolution in the long run. Resolution requires a new relationship to be self-sustaining without any form of coercion.

All conflicts end. Even bitter and decade-long disputes involving several wars have been resolved. On the way to resolution, there may be truces or temporary settlements in which coercion is exerted, but the roots of the conflict remain embedded so that any weakening or withdrawal of coercive mechanisms may risk a new flare-up. Thus, for example, in the Franco-German context, the Treaty of Versailles (1919) was a conflict settlement, whereas the Charter of Paris (1990) was a conflict resolution.

Conflict resolution is at the far end of a spectrum of outcomes that begins with a *diktat imposed by a victor*. But losers usually have leeway since victors are often dependent on local cooperation. Very rarely is the complete destruction of the enemy, or genocide, envisaged. A second outcome is a *negotiated one involving compromise* and lengthy negotiations often accompanied by considerable violence. Such a settlement is contingent on the balance of forces, and should these change, the smoldering conflict may again flare up. A *stasis outcome* is one in which actual physical violence has ended and its recurrence is unlikely, but the protracted conflict continues by other means since major issues remain unresolved. Finally, a full *resolution* is one in which all the parties have their values and interests satisfied. Conflict resolution is the most ambitious outcome sought since resolution based on a noncoercive framework is hard, but not impossible, to achieve.

Conflicts occur at all levels and in all social relationships. Each has its special characteristics, but there are elements in common. General statements can be made about conflict resolution, which apply to international, intercommunal, industrial, marital, or other conflict. Many principles and, indeed techniques, may be similar, albeit with due account being paid to the idiosyncrasies of level and type.

In political conflict, broadly defined, we can identify three traditions with differing approaches, conceptions of its origin and modalities, as well as aspirations for ending it. The first is the realist or power politics tradition that lays emphasis on the nature of international relations in an anarchical society. Societal elements exist, but they are weak and relations are close to the Hobbesian characterization of life as nasty, poor, brutish, solitary, and short. This is particularly true of the international system where the only reliable help is self-help due to the absence of sufficiently strongly held shared values to form the basis for a collective security. The ensuing security dilemma confronts actors with a tragic choice between guns or butter. Judgments are required about the capabilities and intentions of potential adversaries. Moreover, if this security dilemma is also linked to the notion that all individuals and groups have a drive to dominate, then there can be no peace in a real sense. There can be a truce, a cold war, order, and stability, but all this is contingent on a balance of forces between those who dominate and those who are dominated. It is an eternal struggle between the status quo and the revisionist powers. Order and stability are the products of a particular balance of forces, and any major change will upset that balance and a new equilibrium will have to be established to reflect those changes, perhaps on the battlefield. This is a classic case of conflict settlement.

The second traditional approach points to the importance of structures, with conflict being the result of incompatible interests built into structures. Structural violence suggests that peace is more than the absence of overt violence. Structural violence occurs where an actor may be a "happy slave," unable even to conceive of behaving as others can or prevented from so doing by prejudice, ignorance, or lack of access to facilities available to others. Such impediments are often not of a legal nature or enforced by overt pressure, but they are found in the mind, history, education, and culture of individuals and the society within which they live.

Not all structures are instruments of structural violence. Self-sufficiency is not necessary where we can benefit from comparative advantage, as long as the "terms of trade" are seen as fair by all parties in full knowledge of all aspects of the relationship, including the structural ones. Then a role differentiation can be legitimized and with it the structure within which the roles are embedded. Different roles do not necessarily imply an unfair relationship dependent on coercion. All may benefit equitably if the differentiation is acceptable to all with their eyes open.

Dyed-in-the-wool structuralists in a Marxist tradition would regard the concepts of legitimacy and evolutionary change as simply a means of throwing sand in the eyes of the masses. They insist that structures with embedded incompatible interests cannot be reformed and must be destroyed. Peace, or conflict resolution, is thus dependent on the destruction of existing exploitative structures and the creation of a revolutionary peace, if necessary by violence. Increasing the degree of confrontation in society is thus a step toward revolution and, thereafter, conflict resolution.

At all levels, conflict resolution is very different in the pluralist tradition since it gives salience to the notion of harmony of interests. Collective security will work if all the actors in a system come together to agree on the rules of behavior and on how to alter the rules to accommodate change. They further accept that if anybody strays from the fold they shall be brought back into it, forcibly if necessary. Underpinning this is the idea that all human beings have a high degree of shared values and that they can pursue these in a rational manner, in the right circumstances.

Rationality in this context is primarily concerned with the opportunity cost, that is, what other desired goals have to be sacrificed in order to achieve a particular goal. Parties in conflict frequently have a tunnel vision in which they become obsessed by one particular goal, paying a very high price in other desired goals to attain it. To see this in context, actors are encouraged to spell out the full range of their perceived goals and then attempt to minimize the opportunity cost between those goals. The next stage is to repeat the exercise, but this time in the context of others, so that in the long run all parties will be in a position where they are maximizing the totality of their values and minimizing the overall opportunity cost. Such an exercise in rational opportunity costing frequently leads to a consideration of basic human needs that must be satisfied if conflict is to be resolved.

Basic human needs include food and water as well as social needs such as security, identity, participation, self-actualization, and esteem. But these needs are not necessarily in short supply. One person's or group's identity is not necessarily at the expense of another's nor is their security, their development, or their self-actualization. They may be seen as zero sum at a particular moment, but conflict resolution as a process seeks to move to a win-win position. This is difficult since there is no panacea and it often requires third-party facilitation. Nevertheless conflicts can be resolved because the things that we all seem to care about are not inherently in short supply.

Article 33 of the United Nations Charter says that the member states "shall, first of all, seek a solution by negotiation, enquiry, mediation, conciliation, arbitration, judicial settlement, resort to regional agencies or arrangements, or other peaceful means of their own choice." Conflict settlements in the realist tradition rely on an uncertain balance of power. A revolutionary peace requires the destruction of exploitative structures and the eradication of structural violence. The pluralist approach to the process of conflict resolution is of a different ilk. In deep-rooted conflicts where there is little or seemingly no value consensus, negotiation between the parties very quickly descends into a rut of accusation and counteraccusation with little understanding or effective communication. Thus, recourse to other measures under Article 33 involving a third party is frequent. If the goal is merely conflict settlement then a biased mediation from a major actor to impose a settlement may suffice. However, if the goal is conflict resolution, it is more likely that the process is one of facilitation, which has a number of characteristics.

Facilitators in the resolution process are nonjudgmental since they do not pronounce guilt or innocence nor are they directive. They do not contend that they know best. They are, however, extremely supportive of all the parties since the values and interests of all must be met. The process is highly participatory with all the veto holders present, no matter what their past. Finally, there is a disempowerment of the parties by emphasizing that they share a joint problem that they can only resolve together. This approach typifies a more general movement at all conflict levels from judgmental decision making to that of a supportive framework. This is particularly apposite in "new international conflicts," where civil wars, transnational conflicts, and fragile and collapsed states are seemingly endemic in struggles of greed and grievance. Wars and conflicts that go beyond the international system require approaches, methods, and institutions that do likewise. Second-track diplomacy, especially facilitation, becomes thereby a major tool for resolution, involving a variety of actors both governmental and nongovernmental.

The process of conflict resolution is frequently the last resort when the parties know that they can neither win nor are likely to lose. They cannot escape from the conflict, but they know it is costing a great deal in various ways. The process of conflict resolution therefore accumulates many of the hard cases. It is brought in when all else has failed and it is trying to achieve the most difficult goal. However, conflict resolution is relevant earlier in the conflict cycle, as a preventive measure. Much thought has been given to this due to the upsurge of international interventionism and the evident need for emancipatory resolution based on social justice and solidarity in a local context.

A. J. R. Groom
University of Kent
Canterbury, United Kingdom

See also Conflicts; Mediation in International Relations; Peace; War and Peace

Further Readings

Burton, J. (1990). *Conflict: Resolution and prevention*. New York: St. Martin's Press.
Jeong, H.-W. (2008). *Understanding conflict and conflict analysis*. London: Sage.
Miall, H. (2007). *Emergent conflict and peaceful change*. Houndsmill, UK: Palgrave Macmillan.
Mitchell, C., & Banks, M. (1996). *Handbook of conflict resolution*. London: Pinter.
Tidwell, A. C. (1998). *Conflict resolved?* London: Pinter.

Conflict Theory

See Conflict Resolution

Conflicts

In everyday English-language usage, *conflicts* refer to (a) serious disagreements, (b) prolonged struggles (e.g., armed clashes resulting in loss of life, or labor disputes), or (c) the clash of incompatible claims or principles, as in conflicts of interest.

In political science, conflicts can be analyzed as the outcomes of a certain kind of relationship, comparing relationships of cooperation and competition with those of conflict. When there is cooperation, humans work together in pursuit of common objectives. When there is competition, they struggle with one another in contests governed by rules that determine who succeeds and that identify the sanctions to be imposed for noncompliance. When there is conflict, either no rules are recognized or any relevant rules are not properly enforced. While rules may be embodied in laws or treaties, there may be dispute about their application. Moral principles may sometimes be regarded as rules of conduct. It should be noted that even in armed conflicts between states, offenders can be punished for war crimes, while the Geneva Conventions specify the ways in which civilians are to be protected in international conflicts.

Though the absence of conflict is not considered newsworthy, it can be just as important to political science as the analysis of violence. Distinguishing the dimensions of conflict makes it possible to uncover underlying principles, such as has been done in game theory. Less abstract lines of analysis, such as field investigations of the strategies pursued by the parties to conflicts, and the options open to them, can be equally illuminating.

Everyday Language

In the newspapers, and in popular discussion, conflicts are divided into kinds or classes of conflict, such as economic, ethnic, national, racial, and religious conflicts. Yet conflicts do not fall into any natural classification. If they are classified, it can only be for a specific, and limited, purpose.

National conflicts have been of particular interest to political scientists; they are commonly associated with the formation of states and the delimitation of their territory. After major wars, state boundaries have been drawn or redrawn at international conferences convened by what used to be called "the Great Powers." Thus, in the history of Europe, two long-running wars were brought to an end in 1648 by the Treaty of Westphalia; this marked the birth of the concept of the nation-state. After the Napoleonic Wars, the Congress of Vienna in 1815 redrew the political map. After World War I, the Paris Conference of 1919 revised that map and gave birth to the League of Nations. After World War II, the three-power conference at Yalta delimited "spheres of influence"; the United Nations (UN) was established, and its Security Council provided a forum for the regulation of interstate conflicts. The UN has made states responsible for the regulation of conflicts arising within areas subject to their jurisdiction.

The then–UN secretary-general, Boutros Boutros-Ghali, in his 1992 report *An Agenda for Peace*, stated,

> Since the creation of the United Nations in 1945, over 100 major conflicts around the world have left some 20 million dead. The United Nations was rendered powerless to deal with many of these crises because of the vetoes—279 of them—cast in the Security Council, which were a vivid expression of the divisions of that period. With the end of the cold war there have been no such vetoes since 31 May 1990, and demands on the United Nations have surged.

International conferences, such as the Paris Conference of 1919, have often changed state boundaries or created new ones without the agreement of the populations concerned. Campaigning for the restoration to a country of territory occupied by coethnics is known as irredentism. One of the challenges to political science is to explain why, after the ending of the Cold War in 1989, Armenia, Croatia, and Serbia engaged in expensive irredentist campaigns while Hungary, Romania, and Russia did not.

The new international order has been unable to resolve some interstate conflicts. For example, the United Kingdom (UK) decided that it would withdraw from India in 1947 and that the territory should be divided to create the new state of Pakistan. It made this decision in advance of any agreement about the course of the boundary between the two states and about whether Kashmir (i.e., what was then the princely state of Jammu and Kashmir) would form part of India or Pakistan. With partition, Muslims and Hindus moved across the new border, rioting occurred elsewhere, and some 2 million men, women, and children lost their lives. The conflict over Kashmir remains unresolved and is a threat to peace in the region.

The following year (1948) saw the creation of the state of Israel and the intensification of the conflict between Israelis and Palestinians; this conflict has also remained unresolved and is a threat to peace in the Middle East.

The Paris Conference led to the recognition of the Kingdom of Serbs, Croats and Slovenes, as an extension of what had earlier been the Kingdom of Serbia. In 1945, it was reconstituted as the Federal Republic of Yugoslavia. The Federation broke up after Croatia and Slovenia seceded in 1991. Disputes have persisted over recognition of the state of Macedonia (and its name) and over the status of Kosovo, traditionally part of Serbia but with a largely Albanian population.

Elsewhere there have been serious intrastate conflicts, sometimes initiated by elements within the government itself, such as the attempted genocide in Rwanda in which, in 1994, many ethnic Hutus attempted to murder all ethnic Tutsis.

The Dimensions of Conflict

It is difficult to classify conflicts because each conflict is unique. A conflict is, by its very nature, political, but it may also have other dimensions.

Consider the case of Northern Ireland, an entity that was created in 1922 when what is now the Republic of Ireland was allowed to leave the UK and become independent. A narrow majority of the population of Northern Ireland identified as Protestant and wished to remain part of the UK. A minority, identified as Catholic and Republican, believed that no boundary should have been drawn, because the island was one unit. This was not a conflict between two homogeneous blocs, for many Republicans, despite believing that they had legitimate grievances, would have been satisfied without a united Ireland, and many in the majority disassociated themselves from the extreme expressions of Unionism. On both the Unionist and Republican sides, there was more than one political party competing for votes, quite apart from the party that was trying to bridge the divide.

Over the years there were issues that mobilized varying degrees of support either for protest or for the repression of what was perceived as a terrorist movement. Some sort of conflict continued, but it varied in intensity. It had a national dimension because, although the border had been guaranteed by the two states, it was still contested. It had an economic dimension because the Unionists monopolized the better-paid jobs. It had a religious dimension because many identified themselves by the faith they professed. Insofar as there were distinctive communities, it may be said that there was a conflict with an ethnic dimension. Different individuals had different motives for participating in the conflict (or for evading it by emigration). Individual identification with the opposing parties, and with the actions which expressed the conflict, varied continually.

If a broad definition is employed, it could be said that Northern Ireland provides an example of an ethnic conflict, though ethnicity is scarcely its distinguishing characteristic, and a classification as "ethnic" does not contribute anything of value to the analysis of events there.

Pakistanis are affronted that Kashmir, with a majority Muslim population, should remain a part of India, therefore the conflict there also has both a national and a religious dimension. Since the territory is valuable, it has an economic one as well. The conflict between Israelis and Palestinians likewise has a national dimension, reflecting a struggle between, on the one hand, a people with a Zionist state, and, on the other, a people who want either a new state of their own or equality in a reconstituted non-Zionist state. It has a religious dimension and, because of the great inequality of the conditions of the two peoples, an economic dimension as well. What is now the People's Republic of China contends with two conflicts that are either—depending on interpretations of the historical record—within or on its borders. Tibet and Xinjiang have at times been subjected to Chinese overrule; in both, the indigenous people practice their own religions and demand a greater autonomy than that accorded by Chinese policies. Since the parties to each of these four conflicts are distinctive peoples, they are often represented as having an ethnic dimension. This suggests that all persisting conflicts, like all persisting social groups of the kind called communities, are almost certain to be multidimensional.

Theories of Conflict

In the 1960s and 1970s, some sociologists criticized prevailing assumptions that societies were based on consensus and that conflicts were therefore a deviant condition. Such inferences were criticized by Ralf Dahrendorf and others who maintained that, in industrial societies, there were necessarily conflicts of interest, such as those between employees and their employers. Indeed, according to the Marxist version of what was sometimes called "conflict theory," the institutional structure of any capitalist society had to be explained as an imposed scheme for the regulation of class conflict. This was a useful correction of the approach known as structural functionalism, but to call it "conflict theory" was to use the word *theory* to describe a philosophy of society rather than to build a set of testable propositions.

A different example of theory in political science is the application of game theory to the study of conflict. Some conflicts resemble a zero-sum game: If one side makes a gain, the other side loses by an equal amount. Gains and losses sum to zero. Other conflicts have a positive-sum outcome. For example, conflicts serve the common good if they bring underlying disputes to the surface and lead to their resolution. In industrial relations, a threatened strike over pay can force an employer to make changes (e.g., investing in new machinery) that increase production and improve the income of both employers and employees. The sum of gains and losses on both sides is then more than zero. Yet other conflicts either leave both sides worse off, or any gain is more than counterbalanced by losses elsewhere; these then resemble a negative-sum game.

During the Cold War, game theory was used for calculating strategies for military opposition between East and West. The planners reasoned, "If we do this, they may retaliate with Policy A or Policy B. What is then the best course for us to follow in either event?" This was an application in which the two superpowers were seen as playing a game. The same reasoning can picture individuals as players—for example, "Will my gain exceed my losses to a greater degree if I support one side or the other in a given conflict? Or will it pay me to support neither?" This reasoning relies on rational choice theory to calculate the costs incurred by any failure to follow the course that maximizes net advantages.

In recent times, some major conflicts have been resolved by the intervention of a third party. During the Cold War, each of the superpowers intervened to resolve major conflicts within its own sphere of interest lest the opposing superpower exploit them in its interest. (For example, the North Atlantic Treaty Organization [NATO] powers could not allow the conflict between Greece and Turkey over Cyprus to become an armed conflict.) Since the ending of the Cold War, the United States (helped by Canada) has intervened successfully to resolve the conflict in Northern Ireland and, so far unsuccessfully, the conflict over the rights of Palestinians. Superpower intervention is more potent because only such a power is able to promise rewards for all parties and thus hold out the prospect of a positive-sum outcome.

Those who apply game theory liken the parties to a conflict to the players of a game, but they cannot explain why there is a conflict in the first place. In the game of chess, one party plays white and one plays black. Game theory cannot explain which party gets to play white rather than black. Nor has this approach yet been very successful in linking the considerations of domestic policy to those of foreign policy, though models of two-level game theory take a step in this direction. In domestic negotiations, the executive is pictured as absorbing the concerns of interest groups and building coalitions with them; at the international level, the executive then tries to further these concerns without committing to anything that will have deleterious effects at home. When the objectives of actors at both levels are in harmony, an international agreement can be negotiated.

Strategy

That conceptions of interest may be overlaid by other perceptions can be illustrated from studies of conflicts arising within the Hindu caste system of India. That system institutionalized inequality in ways that can be seen as attempting, with only limited success, to prevent competition between caste groups and to make protest illegitimate.

Some conflicts can be observed at the level of the village. A study in the state of Orissa by Frederick George Bailey in the early 1950s described a village in which the population was divided into caste groups (or *jatis*). Almost everyone was a Brahmin, a warrior, a distiller, a herdsman, a potter, a washerman, or a weaver, and there was no exit option. The weavers were "untouchables" (now often referred to as *Dalits*). They knew that the passing of the Temple Entry Act in 1947 had made it an offense to bar Hindus from temples on the grounds of untouchability, and they decided to assert their new right. They notified the local police headquarters that on the occasion of a particular festival, they would, as usual, take their offerings to a temple, but this time they would, like the higher castes, take them into the forechamber. The upper castes mounted a guard to prevent their doing so. The police arrived. They apparently advised the weavers to seek a remedy through the courts and probably said that if there were any more reports of trouble, two constables would be stationed in the village. The upper castes punished the weavers by ceasing to employ them as musicians on festive occasions, but the alleged failure to observe the 1947 Act did not go to court, and the previous equilibrium was restored.

The strategy of the weavers showed calculation in their choice of a particular temple and of a nonviolent form of protest, and in declining the opportunity to pursue legal action. They invited oversight of their dispute, but they did not press for intervention. Everyone knew that if a police unit were to be billeted on the village it would be a collective punishment, a negative-sum outcome. By demonstrating a capacity to press a shared interest, they had presumably gained an enhanced self-respect at the price of a small material loss. They had warned the higher caste groups not to take traditional distinctions for granted.

Interpersonal relations within the village were multidimensional. Throughout, the caste groups cooperated but on terms that were not universally accepted. Dissent over these terms led to the challenge in which all concerned counted the cost of their actions.

Tolerance of what some regard as unjustified inequalities can be changed by the eloquence of mobilizers or by the messages of the mass media. Commentators agree that the highly selective use of television and radio to disseminate misleading accounts of the conduct of members of the national groups in the former Yugoslavia did much to stimulate violent conflict and led to the breakup of the federation. In Rwanda, radio messages inciting Hutus to slaughter Tutsis helped mobilize the mobs responsible for the genocide there. The calculation of strategy at the local level may be skewed by events beyond local control.

Questions and Answers

Political science research on the nature of conflicts needs to delimit the area of inquiry by examining a particular kind of conflict, or a particular aspect of it. Only by delimiting the question posed can an inquiry culminate in the sort of answer that characterizes social science.

Many studies have concentrated on violence, especially deadly violence. As forms of collective violence, riots have been distinguished from feuds, violent protests, terrorism, civil war, and genocide. Some of the questions posed have had a top-down character: Why are there more violent conflicts in some regions than in others? Why does their incidence vary from one historical period or time to another? Why do some localities experience repeated violence? Most of the violent conflict in India is not between caste groups but centers on the hostility between Hindus and Muslims. Why should eight cities, containing just 18% of India's population, have accounted for nearly half of the total deaths from Hindu–Muslim urban violence between 1950 and 1995?

Ashutosh Varshney compared the conflict-prone city of Aligarh with the relatively conflict-free city of Calicut. In both cities there were conflicts between Hindus and Muslims, and in both cities committees to prevent violence had been established. They were effective in Calicut for several reasons, among them the counterbalancing effect of caste divisions among the Hindus. In Calicut, the interest of the political elite lay in the prevention of violence. In Aligarh, by contrast, different sections of the elite could gain from it. The study counted Hindu–Muslim violence as *ethnic* conflict, which is questionable; it took violence as a criterion for determining when an opposition became a conflict, which may not suffice as a criterion of conflict in other circumstances.

Is there an association between the nature of the state and the nature of the violence? For example, it has been argued that whereas a democratic country with a powerful state, such as Japan, experiences low levels of violence, a less democratic regime with a powerful military, such as Turkey, experiences more and that, because of state power, it has to take the form of terrorism rather than of rioting. Such an explanation must be qualified because, in Turkey, the actions described as terrorist have been taken on behalf of a minority seeking recognition as an ethnic or national unit, and there is no comparable minority in Japan.

Some questions have a bottom-up character. Why is one category of persons attacked rather than another? After the assassination of Prime Minister Indira Gandhi in 1984 by a Punjabi Sikh, mobs in Delhi attacked Sikhs, but those they attacked were Sikhs from parts of India other than the Punjab, and the victims were not supporters of Punjabi separatism. In the Northern Nigerian town of Kano, in 1953, and again in 1966, Ibo settlers from the southern regions were attacked but not Yoruba settlers. The Northerners had just as much reason to suspect the political and commercial ambitions of the Yoruba from the Southwest as the Ibos from the Southeast. Why should one group have been victimized while another remained unharmed?

Rioters often justify their actions as responses to provocation, yet sometimes incidents that could be represented as provocative are ignored. Such differences inspire important questions about how violent incidents are organized and the willingness, or ability, of officials (such as police commanders), and of political elites, to suppress them.

Why, it was asked earlier, did Armenia go to war in 1991 with Azerbaijan in order to establish a corridor linking up with an Armenian enclave living in Azerbaijani territory? It proved an expensive venture, and the result has not been acceptable to other states. Why did Croatia try to grab portions of Bosnia inhabited by fellow Croats and Serbia risk so much on behalf of Serbs outside Serbia? Can these actions be attributed to nationalism? If so, why did Hungary not try to renegotiate its boundary with Romania to recover some of the territory it lost in 1920? Romania lost Bessarabia in 1939; later it became the state of Moldova. Why did Romania and Moldova not reunite after 1989? Why, after the dissolution of the USSR, was not more done to bring the 25 million people outside Russia into closer contact with their motherland? A closer analysis of these cases shows, for example, that Armenian nationalism was exceptionally sensitive because of the genocidal persecution of Armenians in Turkey in 1916. Nationalist sentiment in itself achieves little unless there are politicians who can attain power by inciting it, while the voters do not want to carry the costs of incorporating a minority from whom they have grown distant.

Apart from the questions about the policies of states, there are many questions about the behavior, in conflict situations, of the public. Why do otherwise peaceful neighbors kill each other? If they kill others, why do they kill them in one way rather than another? Why are the bodies of victims of collective violence sometimes mutilated? Why are certain sorts of people more inclined to participate in collective violence? Within the same group, some prefer peace and others riot. Why are men more prone to violence than women? Anyone who favors a particular answer to such a question then has to ask why that answer obtains in certain circumstances and not in others.

Conclusion

Germany and France were at war with one another from 1870 to 1871, 1914 to 1918, and 1939 to 1945. In the 21st century, it is inconceivable that either country would declare war on the other because their economies, and the lives of their peoples, are now so closely intertwined. The sources of peace are to be sought in the relations between humans, their understanding of the nature of these relations, and their knowledge of other people. By seeking these sources, political science research aims to advance knowledge about the nature and regulation of conflicts, and it has therefore to account for the absence of conflict as well as for variations in the intensity and scope of conflicts. While the detailed historical examination of particular conflicts is essential, there are general issues that continue to stimulate research.

Michael Banton
University of Bristol
Bristol, United Kingdom

See also Bargaining; Game Theory; Peace Research; Rational Choice; United Nations

Further Readings

Bailey, F. G. (1996). *The civility of indifference: On domesticating ethnicity*. Ithaca, NY: Cornell University Press.

Banton, M. (2000). Ethnic conflict. *Sociology, 34*(3), 481–498.

Dahrendorf, R. (1959). *Class and class conflict in an industrial society.* London: Routledge & Kegan Paul.

Horowitz, D. L. (1985). *Ethnic groups in conflict.* Berkeley: University of California Press.

Horowitz, D. L. (2001). *The deadly ethnic riot*. Berkeley: University of California Press.

Jenne, E. K. (2007). *Ethnic bargaining: The paradox of minority empowerment.* Ithaca, NY: Cornell University Press.

Laitin, D. D. (2007). *Nations, states, and violence.* Oxford, UK: Oxford University Press.

Putnam, R. D. (1988). Diplomacy and domestic politics: The logic of two-level games. *International Organization, 42*(3), 427–460.

Saidemann, S., & Ayres, R. W. (2008). *For kin or country: Xenophobia, nationalism, and war*. New York: Columbia University Press.

Schelling, T. C. (1960). *The strategy of conflict.* Cambridge, MA: Harvard University Press.

Varshney, A. (2002). *Ethnic conflict and civic life: Hindus and Muslims in India*. New Haven, CT: Yale University Press.

Confucianism

Confucianism is the general system of ethics that the Chinese thinker and social philosopher Confucius (ca. 551–479 BCE), his interpreter Mencius (ca. 372–289 BCE), and their early followers advocated to build a moral community of *datong shehui*, known as the Great Harmony Society, in which people could live happy and worthy lives. To build such a community, the Confucian ethical system offered a unique model of benevolent government through moral education and virtuous leadership. This entry first introduces the central values and principles of Confucianism and then highlights its differences from the Western liberal democratic model of government.

Confucian Principles of Good Government

What makes good government? Who should rule and how? What should rulers do to promote good government? On the basis of a positive conception of humans as moral and social beings, not as free and autonomous individuals, Confucius and Mencius sought to address these questions concerning both the ultimate ends and operational means of governance from the perspective of building a harmonious moral community of *datong* (grand harmony), where robbers, thieves, rebels, and traitors had no place, and hence the outer doors remained open and were not shut. Specifically, they conceived of government as an institutional mechanism to achieve such a community. They endorsed economic prosperity, physical security, and popular trust in government as the three essential substantive components of good government.

To create a harmonious and prosperous community, who should rule? In the Confucian conception of good government, the quality of government depends exclusively on the quality of people in the government. It has very little to do with the way in which institutions are organized or how authority is distributed across the institutions (such as, for example, by separation of powers). Accordingly, Confucius and Mencius advocated an open system of meritocracy by dividing people into the two categories of governors and the governed (*The Analects of Confucius*, IV 14). To these and other early Confucians, society becomes harmonious only when those who occupy positions of responsibility are the ones with the ability to discharge those responsibilities.

What specific quality is required of those who should govern? Confucianism endorses the possession of virtue—understanding of and commitment to the common good—as the only proper basis for a claim to governmental authority. Accordingly, Confucius and his followers called for government by the virtuous and wise, not by ordinary people. Yet they did not favor the hereditary system of aristocracy. Instead, they advocated a merit-based system of government, which would allow every citizen to take competitive examinations that test virtues and knowledge. Thus, the most talented would rise to the highest offices. This system of meritocracy replaced the rule by hereditary aristocracy in Confucian Asia long before the same was done in the West.

How should rulers govern the state to build a harmonious and prosperous community? Early Confucians advocated the principle of *minben* as the most important principle of governing. Minben, which originates from the pre-Confucian period, means treating "people" (*min*) as "roots" (*ben*). Thus, they form the foundation of the state. To govern according to this principle is, therefore, to govern for the people (min), for their economic prosperity, and for their physical security just as the roots of trees (ben) should be tended (*Analects*, XIII 9). By embracing ordinary citizens as the roots of government, this principle demands that rulers seek the prosperity and welfare of ordinary citizens as the ultimate end of good government.

To govern according to the principle of minben involves not only providing sufficient food and security but also disseminating virtue throughout society (*Analects*, XII 7). In this regard, Confucius said that people should be educated as soon as their livelihood is secured (*Analects*, XIII 9). Confucianism gives priority to the task of securing livelihood because people can only be expected to behave morally after they are relieved of poverty. As Benjamin Schwartz (1985) points out, Mencius (3a.3) emphasized an economic livelihood as an indispensable precondition for moral education and advocated the hierarchy of human needs as a conceptual tool for governance more than 2 millennia before Abraham Maslow did.

To early Confucians, moral education was essential to the building of a harmonious community because law and punishment were viewed as ineffective means for maintaining order. The higher goal of social harmony could be attained only when people became virtuous and fulfilled their duties to others voluntarily. To build a nation of virtues, therefore, Confucianism emphasized the importance of providing education to all segments of the population. Confucius himself said there is no class distinction in education (*Analects*, VII 7). He was the first scholar to advocate and popularize the notion of mandatory universal education.

The minben principle also holds that good government depends on the mutually beneficial relationship between rulers and the ruled (*Analects*, XII 9). For this reason, Confucius advocated not only good intentions but also benevolent behavior on the part of rulers and the avoidance of cruelty, oppression, injury, and meanness (*Analects*, XX 2). Confucius urged rulers to treat people as generously and honestly as their own guests. They were not to impose on others what they themselves did not desire (*Analects*, XII 2). Only when rulers practice this golden rule of *ren* (benevolence) can the people provide a sense of legitimacy to their government. In other words, rulers can govern people only when they have acquired their trust. Otherwise, people will begin to think they are oppressed (*Analects*, XIX 10).

According to early Confucians, heaven consents to the legitimacy of a government when its people consent to it. According to Mencius (5a.5), heaven sees with the eyes of its people; heaven hears with the ears of its people. They further emphasized popular consent as the most essential component of good government, more essential than sufficient food and sufficient weapons (*Analects*, XX 2). In short, early Confucians proposed an important political theory that holds that governmental legitimacy depends solely on popular support and trust. This notion of governmental legitimacy is at the heart of the contemporary theory of democratic regime consolidation. As Juan Linz and Alfred Stepan (1996) observe, democracy becomes legitimized only when people embrace it unconditionally.

To obtain good government, the early Confucians advocated the use of moral persuasion rather than universally applicable penal laws, because the latter often require coercion. Having assumed that people are inherently good and thus capable of learning and achieving moral development, they can be persuaded to overcome their selfishness and exercise moral discipline instead of being forced to do so by law. By motivating them to follow positive examples and fulfill their duties voluntarily, the rule of virtue or rule by morality can win their loyalty to the government.

Rule by law, on the other hand, involves forcing people to submit to government through punishments, thus making them distrust it. For another reason, early Confucians preferred the rule of morality and customs to that of laws and regulations. They believed that a sense of losing face and being shamed would serve as a more effective deterrent to crime than the sense of guilt that the latter engenders.

The third Confucian principle of good government is known as *zhongyong,* or the doctrine of the Mean. This tells rulers how to make decisions in order to achieve the state of harmony among those with different preferences and interests. To underline the importance of this decision-making principle, Confucius said that perfect is the virtue that is according to the Mean. The core idea of this principle is that going beyond the limit is as bad as falling short (A. T. Nuyen, 2000). It requires being reasonable and moderate without going to the extreme or being one-sided. As a decision-making rule, this principle predates the contemporary median voter theory, holding that politicians should commit to the median policy position preferred by the electorate in order to maximize their votes (Anthony Downs, 1957).

Conclusion

Confucianism equates good government primarily with a benevolent government built on the consent of the people. It also equates good government with primary reliance on virtue and morality. The Confucian governing model calls for government for the people by a virtuous and meritocratic leadership, not government by the people. The Confucian model contrasts sharply with the liberal democratic notion of government, with its emphasis on a government elected by the people and operated according to the rule of law. Due to these differences, there is contemporary debate among scholars such as David Hall and Roger Ames, Samuel Huntington, and Fareed Zakaria over whether China and other East Asian societies shaped by Confucianism can fully accommodate to liberal democracy as practiced in the West.

Nonetheless, we should note that Confucianism has played a vital role in the process of socioeconomic modernization in Asian societies by inducing their governments to promote universal education as a means of national development. To the scholarly community, it offers the original notion of a human-needs hierarchy and theories of governmental decision making and legitimacy. The rich Confucian cultural heritage, moreover, provides the mass publics of the Asian nations the essential premises on which they can construct a viable balance of individual rights and interests with communal goods and obligations as an alternative to the liberal democracy model practiced in the West.

Doh Chull Shin
University of Missouri
Columbia, Missouri, United States

See also Authoritarian Regimes; Elitism; Governance, Good; Legitimacy; Virtue

Further Readings

Anonymous. (1970). *Mencius* (D. C. Lau, Ed. & Trans.). London: Penguin.

Bell, D. A. (2006). *Beyond liberal democracy: Political thinking for an East Asian context*. Princeton, NJ: Princeton University Press.

Confucius. (1979). *The analects of Confucius* (D. C. Lau, Ed. & Trans.). New York: Penguin. (Original work published 551–479 BCE)

De Bary, W. T. (Ed.). (1960). *The sources of Chinese tradition* (Vol. 1). New York: Columbia University Press.

Downs, A. (1957). *An economic theory of democracy.* New York: HarperCollins.

Hall, D., & Ames, R. T. (1998). *The democracy of the dead: Confucius, Dewey, and the hope of democracy for China*. Chicago: Open Court.

Huntington, S. P. (1992). *The third wave: Democratization in the late twentieth century.* Norman: University of Oklahoma Press.

Linz, J., & Stepan, A. (1996). *Problems of democratic transition and consolidation*. Baltimore: Johns Hopkins University Press.

Maslow, A. (1943). A theory of human motivation. *Psychological Bulletin, 50,* 370–396.
Munro, D. (1969). *The concept of man in early China.* Ann Arbor: University of Michigan.
Nuyen, A. T. (2000). Confucianism, the idea of Min-Pen, and democracy. *Copenhagen Journal of Asian Studies, 14,* 130–151.
Schwartz, B. (1985). *The world of thought in ancient China.* Cambridge, MA: Harvard University Press.
Tan, S.-H. (2003). *Confucian democracy: A Deweyan reconstruction.* Albany: State University of New York Press.
Yao, X. (2000). *An introduction to Confucianism.* Cambridge, UK: Cambridge University Press.
Zakaria, F. (2003). *The future of freedom: Illiberal democracy at home and abroad*. New York: W. W. Norton.

CONSENSUS

Unlike many concepts in political science whose meaning is quite contentious—for example, democracy, ethnicity, and state—the concept of *consensus* is among the simplest and most transparent in the discipline. It is universally agreed that the term denotes agreement or unanimity. It is often used in phrases and in conjunction with qualifiers, such as near-consensus, consensus building, international consensus, moral consensus, consensus-based approach, permissive consensus, and cross-party consensus. Although there is strong agreement on the meaning of consensus, there are many concepts that refer to its absence, including cleavage, conflict, competition, contestation, division, majoritarianism, pluralism, and polarization.

The concept of consensus has played a key role in four areas of research within political science: political philosophy, where John Rawls proposed the term *overlapping consensus* to describe a framework for reconciling fundamentally divergent comprehensive worldviews; comparative politics, where Arend Lijphart proposed the term *consensus democracy* as an ideal type within the genus of democratic regimes; economic policy, where the concept of the Washington Consensus has been highly influential since the 1980s; and organization theory, where the concept of domain consensus denotes the fact that different organizational settings are characterized by distinctive norms and behavior. While the concept of consensus has been centrally important in these four significant research areas, it has been used in each without recognition of its key role in the other three research programs. This entry first describes the four uses of the term and then identifies some applications of consensus in political practice.

Four Scholarly Uses of the Concept of Consensus

Overlapping Consensus

Rawls proposed the term *overlapping consensus* to provide a practical solution to a fundamental political problem for liberal democratic theory: How is political order maintained when reasonable people hold deeply divergent comprehensive worldviews? Rawls believed that there is no philosophical solution to the problem, that it is impossible to reach a consensus by rational discussion on the meaning of the public good. His solution was to propose a liberal political framework, one maintained by overlapping consensus among those with fundamentally different comprehensive worldviews. The framework reflects a "thin" theory of the good—that is, one not based on a single conception of the good. However, Rawls claims that this liberal framework of overlapping consensus is not a second-best pragmatic compromise, a mere result of a stalemate among competing worldviews, and a regrettable necessity given that no one worldview is sufficiently powerful to dominate all others. Instead, he claimed that a liberal framework is the best political solution, one that can be a crucial element in competing comprehensive worldviews. (Some critics questioned his optimism on this score.)

Other political theorists, including Charles Taylor and Jürgen Habermas, have advocated seeking to develop consensus through rational discussion, and the concept known as deliberative democracy has highlighted the value of promoting extended discussion to reach consensus rather than foreclosing deliberation by majority vote. (Political theorists who have contributed to the project are James S. Fishkin and Jon Elster.)

From another vantage point, some political theorists have challenged the normative value of consensus. They claim that a hallmark of democracy is divergence, debate, and conflict. For example, Chantal Mouffe praises the value of what she terms *antagonistic pluralism*. At an extreme, Isaiah Berlin warned decades ago of the danger of a totalitarian temptation implicit in seeking and imposing a single consensual understanding of the common good.

Consensus Democracy

Lijphart suggested the term *consensus democracy* as an extension and partial revision of the earlier concept that he developed: consociational democracy. Consociationalism is a pattern of governance that Lijphart first identified in the Netherlands and later applied, in both descriptive and normative terms, to other deeply divided societies. It involves a pattern of decision making in which elites from separate sectors or pillars reach accommodation or consensus across ethnic, religious, or ideological divides.

Lijphart identified consensus democracy as a broader form of political accommodation. It is an ideal type of regime, one characterized by inclusion, in contrast to its polar opposite, what Lijphart termed *majoritarian democracy*. Some key institutional features of consensus democracies are bicameral legislatures, the use of proportional representation to elect the lower house of the legislature, multiparty systems, multiparty governing coalitions that typically promote supermajority support for the passage of significant reforms, and ample protection for minorities. Consensus democracies are associated with broad governing coalitions. Rather than reaching decisions by narrow majorities, consensus democracies promote policy changes only if they enjoy broad political support.

Critics of Lijphart have challenged what they regard as antidemocratic features of his approach. They charge that the accommodations reached by elites in consociational democracies come at the expense of democratic participation and decision making. And requiring supermajorities within consensus democracies enables minorities to block the majority will. Lijphart responded that comparisons between the two regime types suggest that consensus democracies outperform majoritarian democracies on significant measures of good governance.

Washington Consensus

The term *Washington Consensus* was coined in 1990 to describe a policy agenda developed by officials in the U.S. government and the World Bank and International Monetary Fund (IMF), influential international financial institutions whose headquarters are in Washington, D.C. Leaders of these institutions agreed that neoliberal economic policies were the most appropriate and desirable to promote economic development and growth. The Washington Consensus accorded priority to economic "fundamentals," including balanced budgets, price stability, tax and spending cuts, trade liberalization, deregulation, and privatization of state-provided goods and services such as utilities, transport, and education. These goals were to be achieved by adopting structural adjustment policies consistent with neoliberal doctrine. In simplest terms, the Washington Consensus advocated more markets and less state. It was crafted in reaction to the approach stressing state-directed economic policies, which was practiced by less developed countries, especially in Latin America, in the 1950s and 1960s. A key element of the statist approach involved import substitute industrialization—that is, state intervention to foster domestic industrial development, especially by substituting local production of consumer goods for foreign imports. The Washington Consensus was extremely influential for several decades. It informed the aid policies of the U.S. government, the World Bank, and the IMF and shaped the orientation of many governments throughout the less developed world.

The Washington Consensus was far from commanding a consensus. Opponents included progressive economists in the advanced industrial world and governmental leaders and popular movements in countries such as Malaysia, Russia, and Argentina, which experienced economic dislocations and crises after adopting the policies of the Washington Consensus. Critics charged that the Washington Consensus promoted austerity for the broad mass of the population without generating adequate economic benefits, reflected a one-size-fits-all approach that ignored extensive variations among countries, emphasized the value of large industrialization projects that often caused social dislocations and environmental degradation, bypassed democratic decision making, produced financial panics and economic crises, and neglected problems of poverty, corruption, and autocracy.

By the 1990s, modifications occurred within the Washington Consensus. For example, the World

Bank reduced support for large-scale industrial projects such as giant dams; placed priority on developing honest, effective political institutions; and provided for wider consultation in developing policy recommendations. However, as a general matter, changes have occurred *within* the Washington Consensus; no comparably broad alternative has *replaced* the paradigm.

Domain Consensus

The concept of domain consensus has been an important element within the field of organization theory. It refers to the fact that expectations about what is appropriate behavior varies for members of different domains, including sectors such as civil society, the market, and the state; different regions and localities; and specific organizations. The concept refers to expectations about how members in the domain will interact with each other and with nonmembers, about the nature of their respective roles and responsibilities, and about what the collectivity will and will not do.

Political Applications of Consensus

Consensus-based approaches to decision making are relatively rare in nation-states. Bureaucracies usually make decisions by fiat, often after top-down consultation. Legislative decision making is usually made on the basis of majority rule, although in many cases (especially in Lijphart's consensus democracies) supermajorities are the rule. The situation is different in most international organizations, where decisions typically require a consensus of member states. Examples include the Association of Southeast Asian Nations (ASEAN), Asia-Pacific Economic Cooperation (APEC), Southern Common Market Agreement (Mercado Común del Sur, MERCOSUR), and World Trade Organization (WTO). There are exceptions. In the European Union, voting on many issues is by a simple majority, decisions on others is by a weighted majority in which certain member states have plural votes, and some issues are of such importance (notably, the admission of new members or changes of constitutional importance) that unanimity by member states is required. The United Nations Security Council practices a form of supermajority voting. Not only must a proposal pass by majority vote but all five permanent members of the Council must either support or at least refrain from vetoing proposals.

There are many examples of microlevel institutions that employ a consensus-based approach based on the principles of deliberative democracy, including consensus conferences in Denmark, citizens' juries in the United States, planning cells in Germany, and citizens' panels in Britain. These mechanisms often bring together technical experts, government officials, and rank-and-file citizens to analyze issues such as the siting of waste treatment plants or the risks involved in genetically modified organisms and other technological innovations. The goal is to encourage those of diverse perspectives to engage in extended dialogue in order to forge a consensus or at least avoid polarization.

As is evident from this overview, if the definition of consensus is unproblematic, there is no consensus on how the concept is applied and evaluated. It has played a role in highly diverse political analyses and practices. This both testifies to its utility and invites further reflection and experimentation by scholars, practitioners, and citizens.

Mark Kesselman
Columbia University
New York City, New York, United States

See also Cleavages, Social and Political; Competition, Political; Conflicts; Pluralism

Further Readings

Lijphart, A. (1984). *Democracies: Patterns of majoritarian and consensus government in twenty-one countries.* New Haven, CT: Yale University Press.

Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty-six countries.* New Haven, CT: Yale University Press.

Rawls, J. (1993). Lecture IV: The idea of an overlapping consensus. In *Political liberalism* (pp. 133–172). New York: Columbia University Press.

Stiglitz, J. P. (2002). *Globalization and its discontents.* New York: W. W. Norton.

Weisbroad, M. R. (1978). *Organizational diagnosis: A handbook of theory and method.* New York: Perseus.

Williamson, J. (1990). What Washington means by policy reform. In J. Williamson (Ed.), *Latin American adjustment: How much has happened?* (pp. 7–20). Washington, DC: Institute for International Economics.

Conservatism

Conservatism is one of the basic currents of democratic political life, which is founded on a (moderate) critique of liberalism. Conservatism generally accepts the political institutions of liberal democracy, but it denounces the dangers that come from the decline of traditions and from the modern emancipation of individuals. The study of the diverse traditions of conservatism must start with the explanation of its complex nature as philosophy and as political attitude; it requires a historical study of the growth of conservatism from the Enlightenment to the birth of modern democracy through the democratic revolutions of Europe and America and, in the end, a philosophical examination of its theoretical foundations. This entry begins with a definition of conservatism as both a philosophy and a political ideology, briefly traces its historical origins and development in British politics, and concludes with a look at contemporary conservative thought.

Conservatism is both a political philosophy, with its concepts, traditions and classical authors, and a political ideology, which is one of the main currents of political life in modern democratic politics. As a philosophy, it is founded on a critique of modern liberalism, which, as Philippe Bénéton has shown, is directed against what conservatism regards as three great illusions that are the consequences of individual liberalism. The first of these illusions is epistemological and comes from one of the main tendencies of modern philosophy, originating with René Descartes and the age of Enlightenment: It consists in identifying individual reason as the judge of truth. The second one, more directly political, makes individual reason the only foundation of political legitimacy. The third could be named "sociological" and leads to the notion of society as a mere aggregate of individuals, to the detriment of community and/or of social hierarchy. To these illusions, conservatism objects that the individual is never really independent, that all his or her accomplishments have some underlying communitarian or traditional condition that escapes individual reason, and that the neglect of such a debt is a danger for the community and finally even for the individual.

Conservatism is not only a philosophy but also a political attitude that interacts with other currents in political life. To understand its nature, we can start with the usual tripartition that makes conservatism one of the three basic political currents, together with liberalism and socialism: Liberalism is the center of gravity of modern democratic politics but, as it cannot satisfy all the expectations that arise in an open society, it engenders permanent criticisms that can be "socialist," if they argue for a new reconstruction of society, or "conservative," if they try to maintain or restore something of the "world we have lost." In both definitions, conservatism appears not only as an expression of permanent dispositions of human nature but also as a paradoxical product of modernity, which is dependent on what it contests, namely, the faith in individual reason as the foundation of authority, and the philosophical and political order, and which can be defined by three essential features:

1. Conservatism is a *modern criticism of modernity*, which criticizes or even condemns some principles of modernity but does not reject all the consequences of liberal modernity.
2. Even if conservatism was strengthened by its opposition to the French Revolution, its main themes were born before it, within liberal politics, as an internal criticism of classical liberalism.
3. The most central idea of conservative thought is that of the risk of self-destruction of liberal traditions and institutions, which is supposed to be the consequence of a violent rupture with the past and/or with the "natural" basis of human societies. The diverse varieties of conservatism differ in their interpretations of this common idea.

Thus, comprehension of conservatism requires a historical examination of the growth of conservative politics between the 17th and 19th centuries and then a philosophical examination of the diverse expressions of the conservative arguments.

The British Sources of Conservatism

The distinction between liberals and conservatives emerged within English politics, with the opposition between the Tory and Whig Parties, and it took on a new significance with the French Revolution.

Whig and *Tory* are two English words that refer to the two great tendencies—respectively, liberal and conservative—of British politics before the rise of the Labour Party. A Whig is not only a liberal but also a conscious heir of the constitutional English tradition, that of rule of law and parliamentary government; a Tory is devoted to a certain type of social relations, in which the authority of "gentlemen" goes with a protective function and which is connected with High Church Anglicanism. Before 1688, the appellation "Tory" designated the supporters of James II (who the "Whigs" mistrusted because of his alleged "papism" and his commitment to a French-style "absolutist" monarchy), but after the Glorious Revolution the signification of the opposition changed. Most Tories abandoned the absolutist doctrine and partially accepted the Whig interpretation of the English regime as a limited monarchy, but they remained faithful to the institution of prerogative, which authorizes the king to make exceptions to the law of parliament in some special cases, and, more generally, they were devoted to everything that, in the English regime, was favorable to traditional hierarchies. The Tory Party was marginal during the long period of Whig supremacy, which begins with the accession of George I to the throne (1714), but the opposition Whig and Tory became relevant again at the end of 18th century. The Tory Party found a new vigor with William Pitt the Younger, who founded a new alliance between gentry and the commercial elite, while the New Whigs became more democratic thanks to the leadership of James Fox. The French Revolution showed the significance of this evolution; it caused the rupture between James Fox and Edmund Burke who, although a Whig, denounced the French experience as a radical subversion of a tradition that English and even American Revolutions had respected. Whiggism and, more generally, liberalism divided into two tendencies, "radical" and "conservative," which never disappeared afterward.

Has English conservatism experienced a general significance beyond this historical context? One may think so, if one leans on the interpretation of English politics that David Hume gave and on a reflection on the philosophical implications Edmund Burke's criticism of the French Revolution.

David Hume (1711–1776) was certainly a "liberal" with some "conservative" elements in his doctrine, but he was neither a Whig nor a Tory, because he interpreted the English constitution in a sense that explains the permanent division of English politics between successive pairs of parties (first Tory/Whig, then Court/Country). According to Hume, the English regime is both republican, because liberty is expressed through the institution of Parliament, and monarchic, because of the executive power of the king. So the "parties of Great Britain" are for him the expressions of the dual nature of the constitution. In political philosophy, this opposition gave birth to two conflicting doctrines. The Whigs consider that all legitimate government must be founded on explicit consent ("original contract"); the Tories offer an extremist interpretation of loyalism and argue for the idea of "passive obedience" to legal authorities (especially to the king). Hume refuses both principles: "Original contract" cannot explain the real nature of political obligation, and "passive obedience" is incompatible with the real interest of society. Nevertheless, he considers that these two conflicting principles have some kind of truth in the English context, as expressions of the two elements of the constitution of Great Britain, and he shows how some strong "conservative" feeling can be a functional feature of a free government. But, even if he is not himself a "Tory," Hume provides a political philosophy that is "conservative" in another sense. Hume's doctrine of politics accepts the *effects* of revolution without assuming its *principles* because he refuses the illusion of a rational, deliberate reconstruction of political order; he argues for a sort of synthesis between "liberalism" and "conservatism," according to which the durable foundation of liberty requires the stability of institutions and a historical continuity that makes possible a translation of political innovations into the language of tradition. Beyond the Tory pastoral, another kind of conservatism is possible, philosophically skeptical and open to political innovation, intended as a paradoxical mean of conservation of social order (the well-known sentence of novelist Giuseppe di Lampedusa's hero in *The Leopard*—"Everything must change so that everything can stay the same"—could just as well be from Hume).

The French Revolution was seen as the critical turning point in the history of liberalism, which produced a division between two wings: the radical or democratic and the conservative. The most important author is Edmund Burke, who was the first great critic of the French Revolution, as early as 1790, with his major book *Reflections on the Revolution in France*. Burke was himself a genuine Whig, who had supported every great liberal cause, from the fight against political corruption to religious freedom, and who had even accepted American Independence; so his opposition to the French Revolution was especially significant as an expression of the divisions of liberal opinion.

According to Burke, there is an essential difference between the French Revolution of 1789 and the Glorious Revolution of 1688 or even the American Revolution. In one case, people make an exception to the constitutional rule in order to save a regime at risk (the Glorious Revolution of 1688) or to resist tyranny (the American Revolution of 1776); in the other one, a rule is made of the exception itself, by the affirmation that political liberty and the capacity to change the rulers are human rights. For that reason, the French theory of *droits de l'homme* (human rights) represents an aggression against the European political order and a rupture of tradition. On this foundation, Burke builds a coherent conservative doctrine, the influence of which is very large. The basic idea is that of the fragility of human nature, which implies that freedom cannot be natural but emerges only as a product of history and chance, much more than of deliberate action (such is the meaning of the well-known opposition between "rights of man" and "rights of English people"). For this reason, the first condition of liberty is historical continuity, and it explains the ambivalence of the Enlightenment: If, as in England, modernity respects tradition, its effects can be beneficent, but when modern reason revolts against the Christian and aristocratic sources of liberty, it becomes destructive. Thus, the English regime is an example of moderate liberty, in which open aristocracy associates itself with new elites thanks to a transaction between "landed interest" and "monied interest," whereas France is the theater of a mortal conflict between noblesse and bourgeoisie, which produces the insurrection of the mob. This philosophy implies a sound criticism of modern individualism, which constitutes the central argument of Burke in favor of "prejudices" and tradition, which are more rational than individual reason. Burke reasserts many liberal views, most of which come from English and Scottish traditions, but he reinterprets them in a new sense: The market is a way not only of cooperation but also of subordination; the social contract is not equalitarian but is founded on the alliance between God and His creatures. So this theory, which, in Burke's works, remains liberal and, in some way, faithful to moderate Enlightenment, is also one of the main sources of illiberal doctrines, as those of the French Counter-Revolution or German Political Romanticism.

If the legacy of classical conservatism ("Tory," "Humean," or "Burkean") is rich and even imposing in the English-speaking world, it is commonly seen as less impressive in continental Europe and especially in Germany and in France (even if politicians like Guizot and Bismarck or de Gaulle and Adenauer can be considered as great conservative statesmen). In this part of the world, as political liberty comes principally from the French Revolution, liberalism is not traditional, and the criticism of liberal and/or radical politics is often more *reactionary* than *conservative.*

Main Currents of Conservative Thought

To identify the main themes of conservative thought, one can either use a historical investigation (which presupposes some hypothesis about the nature of conservatism) or try to give a general analysis that follows from a view of the uses and functions of conservatism in political life.

The first way can be exemplified by Russell Kirk's *The Conservative Mind*, which is itself a classical work of conservative philosophy. Kirk presented himself as the American champion of a forgotten tradition, who tried to make Americans conscious of the conservative elements of their own culture. He goes back to the origins of American tradition and tries to resuscitate the classical or traditional dimension of the American regime, which he interprets as a mixed constitution, including, thanks to John Adams, some aristocratic features not reducible to the democratic spirit of Jefferson. Referring to diverse American and English thinkers, Kirk proposes a kind of ideal type of conservatism, which combines features such as the belief in a transcendent order, affection for the "variety and mystery" of human existence, acceptance of social hierarchy, and devotion to custom and tradition as conditions of legitimate innovations. Invoking sometimes tradition, sometimes divine revelation or natural law, Kirk is a religious (Christian) and traditionalist thinker whose philosophy conveys a kind of American Toryism.

One can find a good example of a structural analysis of conservatism in Albert O. Hirschman's *The Rhetoric of Reaction: Perversity, Futility, Jeopardy*, which describes the general features of conservative rhetoric when confronted by various examples of liberal or radical arguments or action. Conservatism (which could be called here a "reaction") is supposed to underlie the typical arguments that can be opposed to any attempts to transform or ameliorate the social order. One can say the projected action is dangerous, because it can produce unintended and bad or "perverse" effects (the argument of "perversity"). One can insist on the uselessness of reforms or revolutions, which never can change the human condition or the basic social relations ("futility"). One can, finally, assert that attractive innovations are dangerous for social goods ("jeopardy" vs., for example, the criticism of the welfare state as a danger for civil and political liberty).

Both theories present some interest as descriptions of the conservative spirit or attitude, but to some observers, neither of them seems really sufficient. One possible weakness of Kirk's account is a lack of emphasis on certain themes that are independent of religious faith: Many "conservative" authors, notably in what Robert Nisbet calls the "sociological tradition," were in fact nonbelievers or even atheists, whose criticism of modernity is not founded on nostalgia for the age of faith or on the idea of natural law. As for Hirschman, his model can be applied to any argumentation against deliberate action: the so-called reactionary rhetoric can be used by liberals or progressives (if they want to prevent some "reactionary" decision) and, on the contrary, conservatives can abandon and even reject it when they try to convince people to act with courage or to accept some risks for the public good. There are probably several sorts of conservatism, insofar as the criticism of modernity can be made from different points of view. Perhaps the definition of conservatism must be essentially negative, starting with what conservatives *refuse* rather than with what they *desire*. If one accepts the hypothesis with which this entry starts (conservatism is a basically a modern criticism of modernity), one can say that the main theme of conservatism lies in the refusal of two liberal and progressive ideas: that of the self-sufficiency of the individual and that of a possible social solution to the permanent problems of humanity. This refusal is evidently central in the thinkers that Kirk admires, but it is no less important in other traditions that are not committed to the nostalgia for religious tradition. In fact, besides lovers of transcendence, we find among conservatives many sober (and sometimes even cynical) critics of any sort of idealism. This is notably the case, in the sociological tradition, of the so-called Machiavellians, who contributed notably to the growth and vitality of conservative thought.

One of the most interesting contributions of social science to conservative thought can be found in Vilfredo Pareto's *Treatise on General Sociology*, which is at the same time very critical of humanitarian values and almost skeptical toward traditional ideas. The basic thesis of the *Treatise* is that of the eternity of the division of society between the elite and "the people": It comes from Machiavelli but without any of the democratic connotations that one can find in the Florentine philosopher. For Pareto, society is always divided between the mass of individuals and the ruling, honored (and generally propertied) minority, whose definition is itself value neutral. The elites are not necessarily virtuous or even beneficent: They include all the persons or groups that are in some way above the average. Concerning the *ruling* elite, it is defined by some capacity for domination, either by strength or by cunning (see Machiavelli's division of rulers into "lions" and "foxes"). So this "sociology" is supposed to refute the democratic and socialist idea of an egalitarian society, and it implies some "conservative" or "reactionary" argument in the sense of Hirschman's theory: If division of society is eternal, social change is an illusion, and revolutions are, at best, futile. Besides, Pareto gives not only an *argument* against some specific action but also a general theory, which includes a pessimistic view of human capacity to rationalize society. However, this theory is not "conservative" in Kirk's sense of the word; the social order has no transcendent basis or value, and the first aim of the ruling class is not to promote the social good but to make its domination legitimate. So we can consider Pareto's theory as an example of a conservative theory that abandons the arguments favoring tradition and legitimates the ruling elites without accepting their opinions or arguments.

From this point of view, the most important elements of Pareto's theory are probably, on the one hand, the idea of the circulation of elites and, on the other hand, the division of motives of actions between *residues* and *derivations*. The theory of circulation of elites derives from the elite–mass division: If that division is eternal, all that happens can be only a replacement of an old, exhausted elite by a new and vigorous one. But it implies also some political teachings for the ruling class. If it wants to keep its position, the elite must not submit to humanitarian illusion of egalitarian relations with the ruled, but it must also be open to some elements from the ruled, both to *attenuate* the conflict with other classes and to *regenerate* itself thanks to the incorporation of talented people. So, even if it includes some emphatic exaltation of the elite, the Paretian ideal of a ruling class is close to the moderate English idea of open aristocracy. The distinction between residues and derivations gives a sociological equivalent of a theory of human nature. For Pareto, the main problem of sociology is to explain what he calls "nonlogical" actions—that is, actions in which the means are not adapted to the ends, even if the actor believes that they are. People give many arguments for such actions, but, in fact, experience shows that these justifications are less important than the basic feeling, or instinct, that underlies human actions. So one can distinguish between two classes of motives, *a* and *b*. The first class—*a*, that of residues—includes constant and permanent motives of action; the second one—*b*, that of derivations—comes from the "work accomplished by mind in order to explain *a*" and is much more variable because it comes from "fancy" (*Treatise*, sec. 850–851). Pareto enumerates six classes of residues, among which the two first ones have the greatest importance and pertinence for political analysis. The first class is the "instinct of combinations," which underlies invention and innovation and which is the basis of progress or civilization. The second class is "persistence of aggregates," which explains the fondness for past combinations and the mistrust of innovations. Each of them corresponds to distinct types of human beings and the predominance of one class of residues in a society explains much of its culture and of its political trends. "Conservatism" in the vulgar sense evidently has a strong affinity with "persistence of aggregates," but, for Pareto, the conservation of society (and of an elite) requires an association between the two classes of residues. On one side, we have ancient Athens, or modern France, in which "instinct of combination" predominates: These societies are brilliant but unstable. On the other side, one finds Sparta, or 19th-century Prussia, in which "persistence of aggregates" is very important: They can seem stronger, but they lack the capacity of adaptation to change. The middle between these extremes could be ancient Rome or modern England, in which the combination of the two classes is such that innovation does not jeopardize the conservation of society or of political combination. So, the general theory of human motives (or of human nature) confirms the teaching of the analysis of social stratification: The best way for the elite is to accept change and circulation without adoring them and this "dynamic" conservatism is founded on a secular vision without any faith in Natural Law or in religion.

Pareto has many descendants in social science who could be defined as realists and modern conservatives. The most prominent of them is probably Joseph Schumpeter, whose theory of capitalism and of democracy is a touchstone of conservative sociology. In his political theory, Schumpeter shows that democracy is not really the government of the people by the people and for the people but is simply the free competition of elites for power, with the people as arbiter. However, he admits implicitly that people count, since he considers that democracy can undermine capitalism and even engender "socialism." His vision of capitalism includes a major topic of conservatism. Capitalism is defined by its capacity for innovation, but this strength is at the same time its weakness: The danger is that capitalism could itself destroy the cultural conditions of its development.

So conservatism includes a traditionalist criticism of modernization but is not reducible to that. If it has its roots in disillusionment with the effects of modern liberalism, it is possible to designate as "conservative" some other premodern doctrines, when they are founded on a similar skeptical reaction to social change. Inversely, one can consider that conservatism is always difficult in modern democracies, where innovation is generally more appreciated than stability. One example of this difficulty could be the so-called American neoconservatives, who advocated the use of radical means to attain conservative ends.

Philippe Raynaud
Université Panthéon-Assas (Paris 2)
Paris, France

See also Liberalism; Modernization Theory; Socialism; Traditional Rule

Further Readings

Bénéton, P. (1988). *Le conservatisme* [Conservatism]. Paris: Presses Universitaires de France.

Burke, E. (1999). *Reflections on the revolution in France* (Oxford World's Classics). Oxford, UK: Oxford University Press.

Hirschman, A. O. (1991). *The rhetoric of reaction: Perversity, futility, jeopardy*. Cambridge, MA: Belknap Press of Harvard University Press.

Hume, D. (1987). *Essays: Moral, political and literary*. Indianapolis, IN: Liberty Classics.

Irving, K. (1995). *Neo-conservatism: The autobiography of an idea: Selected essays 1949–1995*. New York: Free Press.

Kirk, R. (1953). *The conservative mind: From Burke to Eliot*. Chicago: Henry Regnery.

Nisbet, R. A. (1966). *The sociological tradition*. New York: Basic Books.

Nisbet, R. A. (1986). *Conservatism: Dream and reality*. Minneapolis: University of Minnesota Press.

Pareto, V. (1953). *The mind and society: A treatise on general sociology*. New York: Dover.

Raynaud, P. (1989). Preface. In *Réflexions sur la révolution de France* (French translation of E. Burke's *Reflections on the Revolution in France*]. Paris: Hachette.

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. New York: Harper & Brothers.

Conservative Parties

Parties are organizations that seek to mobilize public support in order to compete for political power. In principle, attempts at support mobilization can be, and not infrequently are, based on appeals other than those rooted in political views (e.g., the promise of patronage benefits). However, where the publics to be mobilized are very large (as in most liberal-democratic regimes), parties cannot avoid having to compete by espousing alternative sets of public policies—in their turn linked to, and informed by, contrasting political principles or ideologies. At their simplest, then, conservative parties can be defined as organizations bringing together people committed to the quest for power in order to advance an agenda of conservatism—which may in turn be defined as a set of political principles arising from and expressing a commitment to the status quo or the status quo ante. Yet on this definition alone, even parties that most political scientists would *not* see as belonging to the category—for example, the Soviet Communist Party after 1917—would qualify. So there is also a historical criterion that must be fulfilled, since the parties that have tended to attract the label are ones that first emerged in some parts of the world about 200 years ago to counter the demands of liberals—here meaning 18th- and 19th-century Europeans committed to four essential tenets: (1) a belief in the ontological priority of individuals as against society and (2) an unlimited potential for human improvement through the application of reason, therefore (3) a normative commitment to individual freedom and the view that (4) government is legitimate only insofar as it rests on the consent of the governed. Conservatives especially opposed, among the political positions stemming from these tenets, those according to which there was neither inherent value in tradition nor any role for hereditary or religious criteria in political institutions or decision making. This entry reviews conservative political principles in more detail, looking at how they have evolved over time. It also considers how the parties espousing them have been organized, who has supported them and how they have performed electorally, and the role they have played in the party systems that they have been part of.

Political Principles

Conservative parties have, generally, been reluctant to call themselves such, preferring more inclusive titles such as "People's Party" and sometimes even eschewing "party" as well. The reason is that "party" and "conservatism" implied that they sought to make a case for a specific point of view or interest, whereas they saw themselves as defending a way of life, society being viewed as an organic whole—and from their point of view, this was not something they ought to have been obliged to argue over in the first place. From this outlook stemmed their professions of antipathy to "ideology"—the product of reason—with its capacity to disarticulate a body politic whose parts were integrated, not by design but thanks to the accumulated wisdom of the ages. This did not mean that the conservative position was one of opposition to change of whatever kind: On the contrary, societies, like organisms, had to adjust to new circumstances if they were to survive. But it meant that change was to be adaptive, not revolutionary.

The role of change in the conservative outlook has meant that over the years its agendas and programs have—paradoxically—altered much more radically than have liberal and socialist agendas and programs with their greater attachment to abiding principles. Of the changes and trends that required this alteration, the two most significant following the conservative parties' emergence were the spread of industrialism and extension of the franchise. The first brought about, in many instances, a gradual reconciliation and integration of the interests of the landed elite with those of the rising bourgeoisie—the successful pursuit of whose activities required priority to be given to the quintessentially liberal principles of due process, the rule of law, and meritocracy. In such instances, the principles became conservative ones as well. The second trend threatened to render conservatism unappealing on account of its traditional opposition to democracy—but also offered the means to its survival. Conservatives' defense of the established order had always entailed a paternalistic sensitivity to the legitimate concerns of ordinary people and the poor. The mass franchise now enabled and required a renewed emphasis on this social reform aspect of the conservative agenda.

A third significant trend was the growth of nationalism. Originally popular and democratic, it was not therefore a distinctively conservative principle. However, as it increasingly became an accomplished fact—through the unification of Italy and Germany and similar events—it offered a means of countering the most significant threat posed by the advent of the mass franchise: the growth of socialist parties and ideologies of international workers' solidarity.

The final major adjustment in conservative thinking took place in the period after World War II. Initially, the collapse of fascism had brought a "technocratic" kind of conservatism. A social-democratic hegemony based around the principles of the welfare state and Keynesian demand management had gone hand in hand with the growth of so-called catchall parties seeking votes wherever they could find them and bending principles and ideology to that end. Under these circumstances, conservatives found the distinctiveness of their outlook increasingly difficult to convey, seeming to many to stand for little more than a claim to be able to manage politics more competently than their rivals. Later, the increasing difficulties encountered by social-democratic politics in the economic sphere, coupled with the growth in the social sphere of new movements pressing the demands of environmentalists, peace activists, feminists, gays, and ethnic minorities, provided conservatives with a renewed opportunity to press a distinctive agenda. This gave the pride of place not only to the old liberal principles of free markets and the minimal state but also, in opposition to the left libertarianism of the new social movements, to the strong state: a state that would enforce free markets through firmness in welfare and trade union matters and robustly defend laws embodying traditional social values.

Organization

In most cases, conservative parties originated as informal groups within predemocratic legislatures, only later, with extensions of the franchise, developing formally organized extraparliamentary structures covering the national territory. Even then, in many instances, the degree of formality of these structures for long remained incomplete and the real power acquired by them limited. The British Conservative Party, for example, did not have individual members before the 20th century, and until the 1970s, its local associations not infrequently refrained from fielding candidates in municipal and county council elections: giving, in such instances, a free run to "Independents," who were Conservatives at the national level. This was a convenient way of ensuring that local influence remained in the hands of preexisting Conservative elites and did not need to be shared with less privileged groups in the party. Nowadays, there is little that is especially distinctive about conservative parties from an organizational point of view. They differ between themselves depending on the mix of resources—money, activists, charismatic leadership, and so on—available to enable them to compete. Like parties generally, their penetration of society has declined in the postwar period, with internal distributions of power increasingly consolidated in the hands of elected officials, progressively more important, with the spread of the media, for mobilizing a mass electorate.

Social Base and Electoral Performance

Generally speaking, the likelihood of voting for a conservative party increases with religious observance and the belief that religion is important, with high socioeconomic status, with age, and with a view of oneself as being located on the right of the left–right spectrum. However, generalizing about the *precise* impact of variables such as these is impossible for two sets of reasons, the first being the influence of context in mediating their impact. For example, if in Britain the Conservatives have traditionally attracted very high levels of support among farmers, the same cannot be said of Conservatives in places such as Norway and Sweden, whose party systems, from the 1920s onward, included influential agrarian parties. Second, there is the problem of the choice of parties to include. Frequently in comparative analyses, conservative parties have been grouped together with other parties of the moderate right, such as Christian democratic parties with which they have sometimes shared important features, for example, size and significance in the party system, ideological flexibility, and the conviction that, in standing for the interests of the entire nation, theirs is more than a political point of view on a par with the rest. The choice makes a significant difference. Comparing the British Conservative Party with the pre-1990s Italian Christian Democrats, for example, would reveal not similarities but enormous differences, both quantitative and qualitative, in the impact of religious observance and left–right self-placement on propensities to vote.

In few other places in the postwar period have conservatives matched the electoral success of the British Conservative Party. Again, much depends on the choice of specific parties to include, with particular controversy surrounding the status of the U.S. Republicans and Ireland's Fianna Fáil, but if we include only the uncontroversial cases, then the British Conservatives' only real rival in Western Europe is Greece's New Democracy (which has averaged 43.9% of the vote in parliamentary elections since 1974 as compared with 37.8% for the Conservatives). Elsewhere in Western Europe, the best average performance has been by Luxembourg's Christian Social People's Party (with an average of 32.3% since 1974). In some countries, conservative parties are essentially nonexistent: the contrast with countries where they are strong being due to differences of party-system context inherited from the past.

Role in Party Systems

In Britain, the arrival of democratic politics enabled the conservatives to consolidate a position as one of two large parties by virtue of the relative insignificance of three of the major social cleavages—that is, center–periphery, state–church, land–industry—which elsewhere spawned more complex party systems; and by virtue of the fact that extensions of the franchise had begun early and taken place over a long period of time, enabling the party to adapt successfully to the emergence of the fourth major cleavage, that between owners and workers.

In other major European countries—Germany and Italy—the party systems supporting the reemergence of democratic politics after World War II could find no place for conservative parties because of the experience of Nazism and fascism: This made it impossible for anyone to espouse important conservative themes such as nation, tradition, order, and hierarchy without arousing painful memories and the suspicion that they harbored extreme right-wing sympathies. Consequently, in these countries the role of principal party of the center-right came to be played by a Christian democratic party.

In France, conservatism, though present, was able to make less headway than in Britain, for the country's deep divisions—between church and state; between monarchists, parliamentarians, and

Bonapartists; between land and industry; and between owners and workers—gave it a highly fragmented party system within which conservatives had many significant rivals. Meanwhile, the localism of party politics and the two-ballot system inhibited the development of cohesive national parties—and, thus, the emergence of conservatism as a unified force. Though the constitutional overhaul that ushered in the Fifth Republic in 1958 brought an initial decline in polarization, a degree of party–system fluidity continued, and today at least two significantly sized entities—the Union for French Democracy and the Union for a Popular Movement—themselves coalitions, can claim to provide a home to conservatives.

At the European level, the main transnational party organization embracing conservative parties is the European People's Party. This also encompasses Christian Democratic parties and the internal life of the organization is not without tensions between pro-European (mainly Christian Democratic) and Eurosceptic (mainly British Conservative) members.

James L. Newell
University of Salford
Salford, United Kingdom

See also Christian Democratic Parties; Conservatism; Electoral Behavior; Liberalism; Nationalism; Party Organization; Party Systems

Further Readings

Burke, E. (1962). *An appeal from the new to the old Whigs*. New York: Bobbs-Merrill. (Original work published 1791)

Layton-Henry, Z. (Ed.). (1982). *Conservative politics in Western Europe*. London: Macmillan.

Norton, P. (Ed.). (1996). *The Conservative Party*. Hemel Hempstead, UK: Prentice Hall.

Scruton, R. (2001). *The meaning of conservatism* (3rd ed.). Basingstoke, UK: Palgrave.

Von Beyme, K. (1985). *Political parties in Western democracies*. Aldershot, UK: Gower.

Constitutional Engineering

Constitutional engineering refers, generally, to the process by which political actors devise higher law, which is usually—but not always—specified in a formal written document and labeled the "constitution." Constitutional design is a close synonym. Any particular instance of constitutional engineering must deal with certain basic questions of organization and process. These include designating who is to be involved, when that involvement takes place, and how the actors are to proceed in formulating, discussing, and approving a text. Although there are conceivably as many variants in the process as there are constitutions, several common patterns emerge. This entry describes the factors that distinguish instances of constitutional engineering and some of the typical patterns.

Stages

Constitution making occurs in discernible stages, some of which resemble an ordinary legislative process familiar to many drafters in consolidated democracies. A schematic design of these phases might include, in sequential order, the mobilization of interests (and counterinterests), drafting, consultation, deliberation, adoption, and ratification. These different stages interact with the possible actors who might fill the roles to create a matrix of options for designers. Afghanistan's constitution of 2004, for example, was drafted in relative secrecy by a commission with foreign advice and then sent to the president's office before deliberation and adoption at an inclusive constituent assembly, the Loya Jirga. In this model—which appears to be relatively common—each stage is potentially consequential, although it is likely that inertial forces and the power of agenda setting will apportion disproportionate influence to actors involved at earlier stages. Still, it is quite possible that early-stage actors will anticipate the preferences and needs of later stage actors, thus mitigating any sequence effects. Jon Elster has introduced the vivid distinction between upstream and downstream constraints in the process: Upstream constraints are imposed by the powers setting up the constitution-drafting body, whereas downstream constraints result from the anticipation of preferences of those involved in later stages. Ratification by public referendum, for example, is a downstream constraint that can hamstring leaders in an earlier stage who recognize that their document must ultimately obtain public approval.

Actors

Perhaps the most critical variable in constitution making has to do with which actors are included in the process. Actors involved in constitution making can include expert commissions, legislative bodies or committees, the executive, the judiciary, national conferences, elite roundtables, transitional legislatures, specially elected constituent assemblies, interest groups and nongovernmental organizations, foreign advisors, and the public itself. Public involvement has become the subject of particular attention in recent years and is urged by scholars, governments, and international organizations. But not all constitutions involve the public, and some are drafted by a handful of leaders behind closed doors.

Certainly, a central dimension on which constitution-making processes differ is the degree of public participation. Because the constitution is the highest level of lawmaking and provides the ultimate rule of recognition for lawmaking processes, it arguably requires the greatest possible level of legitimation in democratic theory. In an ideal world, one might desire universal consent over the rules of society, a standard that is obviously impractical. Higher levels of participation are presumed to function like supermajority rules, restricting the adoption of undesirable institutions and protecting prospective minorities in the democratic processes that are established. Participation thus legitimates and constrains, substituting inclusive processes for consent to make effective government possible.

The modal form of participation in constitutional design is the power to approve the charter, usually by referendum on the final document as a whole. Available data suggest a significant trend, beginning in the early 20th century, toward public ratification. Approval by referendum may be an increasingly popular mode of public involvement, but it is clearly a limited one in that it involves only an up or down vote over a package of provisions. Since at least World War II, however, participation in constitutional design has become more direct and has penetrated more deeply (or at least earlier) in the process. One common approach is to involve the public in selecting those who will draft or deliberate over aspects of the charter. This sort of voice is possible whether the representative group is a constituent assembly elected expressly for the purpose or is a regular legislature that takes on the project in addition to other duties. Some constitutional processes have experimented with more bottom-up methods of direct democracy, such as the citizen initiative, in which ideas can bubble up from civil society.

Still another mode of participation involves direct consultation with the public or representative groups at various stages, which might occur before, during, or after the drafting of the initial text. The drafting phase seems to be especially crucial because we can expect a fair degree of inertia in the later stages of the process. But the phase is also likely to be the least participatory, given the challenges of writing-by-committee, much less writing-by-nation. Indeed, in some well-known cases, the public is excluded from the drafting process and not consulted at all.

Of course, actors and their accompanying constraints may come from outside, as well as inside, a state's borders. An extreme case is that of the "occupation constitution," a document drafted when a country is under the control of a foreign military power. Such constitutions are usually presumed to have less involvement on the part of local actors and hence to be less legitimate. Some scholars believe that international involvement creates disincentives to enforce the constitution locally, as actors will strategically acquiesce to conditions they have no intention of fulfilling simply to remove external oversight.

Of course, international constraints on constitution making can range in their intensity and degree of coordination, from borrowing to imposition. External influence need not be as blatant as in occupation constitutions. Constitutional drafting that occurs concurrently with peace negotiations often attracts international advisers and interests, be they donors, creditors, interested states, or the United Nations. The prospect of future membership in the European Union, for example, led some Eastern European countries to make modifications to their draft constitutions at the behest of the Council of Europe. Many accounts of foreign borrowing point to the decisive role of influential consultants who appear to be part of a cottage industry of constitutional advisors.

The voluminous literature on policy diffusion reminds us that any sort of policy or institutional reform will be a highly interdependent process.

Constitution making—often undertaken during moments of crisis when states are at their most amenable to foreign models and suggestions—may be especially interdependent and networked. Certainly, scholars have long noted a high degree of similarity across documents, and nearly anyone privy to the details of a case of constitution making can recount an episode of international borrowing. The persistence of presidentialism in Latin America, the use of French and Westminster models of government in former colonies, and the recent use of national conferences in Francophone Africa are all examples of diffusion that occurs at a subglobal level. Given the persistent centrality of the U.S. Constitution to the American legal academy, there has been a fair amount of interest in documenting its influence over the years, but other constitutional models have also clearly had some impact.

Other Conditions

There are potentially consequential aspects of process other than the identity of the actors involved. Some constraints reflect the circumstances that lead to constitution making in the first place. The conventional wisdom is that constitution making is coincident with a cataclysmic event of some kind, such as war, coup, economic crisis, or revolution. In fact, some evidence suggests that, although crises do frequently precede constitutional reform, the degree of noncrisis constitution making is probably underestimated. Sweden's 1972 reform of its 163-year-old constitution is a prominent example of crisis-free reform. The various socialist constitutions, such as those in the Soviet Union (1936, 1977) and China (1982), seem to follow the installation of new leaders, a practice that was often justified by the Marxist view of evolution in stages. These different patterns, reflecting various degrees of crisis or continuity, will affect the process, creating an atmosphere of either urgency or deliberation.

The process can also vary in terms of time involved. At one extreme, the secretive process that led to Myanmar's 2008 constitution took 17 years. At the other extreme, a small group of American bureaucrats working for the occupation authorities drafted the basic form of Japan's 1946 constitution in a little over a week, and the entire process, including elections, legislative deliberation, and approval by the emperor, took a mere 8 months. Both of these cases are distant outliers. On average, constitution making takes around 16 months. Anecdotally, constitution-making processes involving either a very short or very long amount of time seem to occur in nondemocracies. Speedy processes do not allow sufficient time for mobilization of the public and civil society, whereas extended processes are unlikely to hold public attention for the duration.

Another dimension on which processes differ is the size of the deliberative body, an issue that has also plagued those who design legislatures. The concern is that large bodies—which have the advantages of minimizing deal making and assuring representativeness—can be unwieldy and lead to collective action problems.

Constitutions and their design have long been central to the work of political scientists since at least Aristotle. For whatever reason, however, the scientific study of the rules that govern the process of constitutional engineering has lagged. This state of affairs is regrettable but, happily, quite remediable. Processes of constitutional design and adoption vary widely along a number of identifiable dimensions. Speculation about the effects of these differences runs rampant and evidence will undoubtedly be close behind.

Zachary Elkins
University of Texas at Austin
Austin, Texas, United States

See also Constitutionalism

Further Readings

Carey, J. M. (2007, April). *Does it matter how a constitution is created?* Paper presented at the symposium on Exporting Democracy: What Democracies Can and Cannot Do, University of Texas at Austin.

Elster, J. (1995). Forces and mechanisms in the constitution-making process. *Duke Law Journal, 45,* 364–396.

Ghai, Y, & Galli, G. (2006). Constitution building processes and democratization. In *Democracy, conflict, and human security: Further readings* (pp. 232–249). Stockholm: Institute for Democracy and Electoral Assistance.

Hart, V. (2003, July). *Democratic constitution making* (Special Report No. 107). Washington, DC: U.S. Institute of Peace. Retrieved October 27, 2010, from http://www.usip.org/publications/democratic-constitution-making

Horowitz, D. L. (2002). Constitutional design: Proposals versus processes. In A. Reynolds (Ed.), *The architecture of democracy* (pp. 15–36). New York: Oxford University Press.

CONSTITUTIONALISM

Constitutionalism is sometimes regarded as a synonym for limited government. On some accounts, this doctrine is associated in turn with minimal or less government. But that is only one interpretation and by no means the most prominent historically. A more representative general definition would be that constitutionalism seeks to prevent arbitrary government. At its most generic level, arbitrariness consists in the capacity of rulers to govern willfully—that is, with complete discretion—and to serve their own interests rather than those of the ruled. Constitutionalism attempts to avoid these dangers by designing mechanisms that determine who can rule, how, and for what purposes. However, constitutional traditions differ as to what precisely counts as an arbitrary act and which mechanisms offer the best defense against their occurring. The classical, neo-republican tradition of political constitutionalism identifies arbitrariness with domination of the ruled by their rulers and seeks to avoid it by establishing a condition of political equality characterized by a balance of power between all the relevant groups and parties within a polity, so that no one can rule without consulting the interests of the ruled. The more modern, liberal tradition identifies arbitrariness with interference with individual rights and seeks to establish protections for them via the separation of powers and a judicially protected constitution.

Both traditions are present within most democracies and can be found side by side in many constitutions. The first tradition focuses on the design and functioning of the democratic process, including the selection of electoral systems and the choice between presidential or parliamentary forms of government, of unitary or federal arrangements, and of unicameralism or bicameralism. Although the detailing of these procedural mechanisms and the relations between them usually forms the bulk of most constitutional documents, their constitutional importance has come to be eclipsed—in legal circles particularly—by the second tradition. This view emphasizes the specification and judicial protection of the different competences of the political system and of constitutionally entrenched rights by a constitutional court. Political theorists and scientists disagree, however, on whether these two traditions are complementary, mutually entailed, or incompatible. The second is often seen as necessary to ensure the fairness of the procedures and/or the outcomes of the first. Yet it lays itself open in turn to doubts that courts are, or could ever be, truly bound by constitutions so that law rather than judges rule and if so whether judicial processes are not more arbitrary and prone to error for deciding constitutional outcomes than the democratic procedures and outcomes they are often thought legitimately to limit. In the following sections, this entry traces these two traditions and then turns to exploring their respective advantages and disadvantages and any tensions and complementarities that exist between them.

Two Traditions of Constitutionalism

Political Constitutionalism: From Mixed Government to Representative Democracy

The theory of mixed government originated with ancient thought and the classification of political systems on the basis of whether one, a few, or many ruled. According to this theory, the three basic types of polity—monarchy, aristocracy, and democracy—were liable to degenerate into tyranny, oligarchy, and anarchy, respectively. This corruption stemmed from the concentration of power in the hands of a single person or group, which created a temptation to its abuse through allowing arbitrary rule. The solution was to ensure moderation and proportion by combining or mixing various types. As a result, the virtues of each form of government, namely, a strong executive, the involvement of the better elements of society, and popular legitimacy, could be obtained without the corresponding vices.

Three elements underlie this classic theory of mixed government. First, arbitrary power was defined as the capacity of one individual or group to dominate another—that is, to possess the ability to rule them without consulting their interests. To be dominated in such an arbitrary way was to be reduced to the condition of a slave who must act as his or her master wills. Overcoming arbitrariness so conceived requires that a condition of political equality exist among all free citizens. Only then will no one person or group be able to think or act as the masters of others. Second, the means to minimize such domination was to ensure none could rule without the support of at least one other individual or body. The aim was to so mix social classes and factions in decision making to ensure that their interests were given equal consideration, with each being forced to "hear the other side." To quote another republican motto, "The price of liberty is eternal vigilance," with each group watching over the others to ensure none dominated them by ignoring their concerns. Third, the balance to be achieved was one that aspired to harmonize different social interests and maintain the stability of the polity, preventing so far as was possible the inevitable degeneration into one of the corrupt forms of government.

Thus, mixed government provides a model of constitutionalism according to the institutions that structure the way decisions are taken. Although elements of the theory can be found in Aristotle's *Politics*, the locus classicus is Book VI of Polybius's *Histories*. He underlined its prime purpose as providing mechanisms whereby no individual, body, or group could rule alone, thereby curbing the descent into tyranny, oligarchy, or anarchy. Polybius regarded the republican constitution of ancient Rome as exemplifying this theory. Thus, the consuls provided the monarchical element, the Senate provided the aristocratic, while the popular element was represented by the Tribunes of the People, the Plebeian Council, and the electoral, judicial, and legislative powers the people could exercise directly. As he noted, the key feature of Roman republican government was that each of these three groups exercised slightly different powers but required the cooperation of the others to do so. So consuls might exercise war powers, yet they needed the Senate to approve generals, reward them, and provide the necessary funds, while the people approved treaties and could try high officials and generals for misconduct. Meanwhile, the more-executive roles possessing the most discretion were further weakened by their power being shared among multiple officeholders and its being dependent on elections and of short duration. Thus, there were two consuls, each able to veto the other's decisions, 10 tribunes with similar countervailing powers, and so on, with none able to hold office for more than a year.

The resulting need for different groups to work together was summarized in the slogan *Senatus Populusque Romanus* (The Senate and the Roman People, frequently abbreviated to SPQR). In reality, though, their relationship was far from harmonious, with the patrician element largely predominating, except when factional disputes led a given group among them to seek the support of the plebeians. The conflict between social classes was given greater emphasis by Machiavelli, who offered a radical version of the Polybian argument in his *Discorsi*. He observed how all polities contain two classes, the nobles (*grandi*) and the people (*popolo*), whose desires conflict. However, he claimed that their discord, far from being destructive, actively promoted "all the laws made in favour of liberty"—each was led to promote freedom by virtue of seeking ways of checking the arbitrary power of the other. However, like Polybius, Machiavelli believed that all systems ultimately became corrupt and degenerated into either tyranny or anarchy—the balance of power merely served to stave off this inevitable cycle.

The 17th and 18th centuries brought three main changes to the doctrine. The first, explored below, was the development of the separation of powers as a variation on the doctrine of mixed government. The theory of mixed government involves no clear distinction between the different branches of government. Executive, legislative, and especially judicial tasks were shared between the different social classes and exercised by all the government bodies. Indeed, the popular element exercised certain legislative and judicial functions directly through plebiscites and as jurors. The second change was in the type of "balance" mixed government was supposed to achieve. The classic theory took the idea of the "body" politic literally. Just as bodily health was said to rely on a sound physical constitution and a balanced diet and way of life, so the health of the polity depended on a sound constitution that achieved a "natural" balance between the various organs and "humors" of the political body. As we saw, in line with this organic imagery, the aim was to hold off the inevitable degeneration and corruption of the system. Balance was a static equilibrium, designed to maintain the status quo. However, the 17th and 18th centuries saw a new, more dynamic notion of balance, inspired by Newtonian physics and based on mechanics and physical forces. In this conception, balance could involve a harnessing of opposed forces, holding them in a dynamic equilibrium that combined and increased their joint power. The change can be seen in the notion of the "balance of trade," which went from being an equal exchange of goods between states to become a competition between trading nations that encouraged their mutual productivity and innovation. In this account, the "cycle of life," where growth was followed by decay, became replaced by the idea of progress, in which change and transformation had positive connotations.

The third development drew on these two. This was the idea that political balance now consisted in the competition between government and a "loyal" opposition. As parties evolved from simple factions and patronage networks among rivals for office, and became electoral machines defined as much by ideology and social composition as by the personal ambitions and interests of the political class, they became the organs of this new type of balance. In keeping with the older theory of mixed government, one of the virtues of parties was their ability to mix different social classes and interests and combine them around a common program. Indeed, just as economic competition led rival firms to compete over price, innovate, and explore untapped markets, so electoral competition led rival parties to compete over policy efficiency and effectiveness, devise novel forms of delivery, and focus on areas appealing to different sections of the electorate. This modern form of political constitutionalism has proven constitutional in both form and substance. Equal votes, majority rule, and competitive party elections offer a mechanism for impartially and equitably weighing and combining the views of millions of citizens about the nature of the public good. And in making politicians popularly accountable, it gives them an incentive to rule in nonarbitrary ways that respond to the concerns of the different minorities that form any working majority, thereby upholding both rights and the public interest rather than their own interests.

Meanwhile, mixed government has developed in new ways through federal and convocational arrangements that likewise seek to ensure that different kinds of interest are involved in the policy- and lawmaking processes on an equal basis. Yet nobody would deny that the systems of most democracies are far from perfect, and it has become increasingly common to look to other constitutional traditions to rectify these problems.

Legal Constitutionalism: From the Separation of Powers to Rights and Judicial Review

According to Article 16 of the French Declaration of the Rights of Man and of the Citizen of 1789, "A society where rights are not secured or the separation of powers established has no constitution at all." Though widely accepted today, this view was novel at the time, shaped by the experience of the English, American, and French revolutions. The separation of powers developed out of the theory of mixed government during the English civil war of the mid-17th century. In 1642, Charles I belatedly invoked the doctrine of mixed government to defend the joint rule of Monarch, Lords, and Commons as implied by the notion that Parliament meant all three (the doctrine of "King in Parliament" as the sovereign body of the realm). His execution posed the problem of how to control government in a society without distinctions of rank. Dividing the executive, legislative, and judicial functions between three distinct agencies appeared to provide a response to this dilemma. However, it took some time to evolve. Although Book 11, Chapter 6 of Montesquieu's *The Spirit of the Laws* has been credited with offering a definitive statement of the doctrine, his account still bore the hallmarks of its origins in the system of mixed government—not least because of its being based on an analysis of the British parliamentary system and the respective roles of monarch, lords, and commons within it. The functional division also remained far from clear-cut, with the judicial branch and function still imperfectly differentiated from the other two. Only with the drafting of the U.S. Constitution and the debates surrounding it, most notably the *Federalist Papers*, did the doctrine emerge in its mature form.

The underlying rationale of this separation is that individuals or groups should not be "judges in their own cause." The division between the three branches aims to ensure that those who formulate the laws are distinct from those entrusted with their interpretation, application, and enforcement. In this way, lawmakers are subject to the same laws and so have an incentive to avoid self-interested legislation and to frame it in general terms that will be equally applicable to all. These laws then guide the decisions of the executive and judiciary, who because they are similarly under the law also have good reason to act in an impartial manner. Separating functions also brings the efficiency gains associated with the division of labor. In particular, the activity of the legislature is made less cumbersome through delegating more short-term decisions to an executive branch capable of acting with greater coherence and dispatch.

On its own, it is unclear how effective this separation is. Not only are the four functions hard to distinguish clearly, but unless a different group operates each branch, there is nothing to prevent their acting in concert. However, four other theoretical developments accompanied the shift from mixed government to the separation of powers that changed its character. First, mixed government had been challenged earlier by theorists of sovereignty, such as Jean Bodin and Thomas Hobbes, who regarded the idea of dividing power as incoherent. The separation of powers came into being in a context shaped by the notion that at some level power had to be concentrated, and, in the context of the English, American, and French revolutions, the natural assumption was to shift the sovereign power of the monarch to the people as a whole. Second, the notion of the people as a whole was likewise new. Previously, the "people" had simply meant the "commons" or the "many." The whole people became the authors of the constitution, which as the embodiment of their will became sovereign over the will of any subdivision of the people, including the majority. Third, as a corollary, constitutions became entrenched written documents expressing a "higher" law, which could be amended only by the people as a whole or by some supermajority that could plausibly be said to represent their will. Fourth, notions of rights became key aspects of the constitution. Initially rights were no more an intrinsic part of the separation of powers than they had been of mixed government. The Bill of Rights was an appendix to the U.S. Constitution, which had previously been confined to describing the system of government. Nevertheless, the securing of individual rights gradually became the goal of all constitutional arrangements.

These four developments, but particularly the last two, had a tremendous impact on constitutionalism and proved crucial in moving it in a legal and especially a judicial direction. Within the "pure" theory of the separation of powers, all three branches were coequal. As with the theory of mixed government, the aim was to prevent any one section of society dominating another by obliging each to collaborate with the others. If anything, the legislative power was logically prior to the others—producing in the U.S. scheme federal and bicameral arrangements within the legislature that harked back to the doctrine of mixed government and a clear division between the legislature and executive. As noted earlier, the distinctiveness of judicial functions was weak in the doctrine of mixed government and slow to emerge in the theory of the separation of powers. However, making a legal document sovereign—only challengeable by the sovereignty of the people as a whole—inevitably empowered the judiciary, particularly given the comparative length of judicial appointments and their relative isolation from electoral pressures by contrast to the other branches. The judiciary now decided the competences of the various branches of government, including their own, and set limits not only to the processes of government but also to its goals with regard to individual rights. These features have come to define modern constitutionalism and are reflected in all the constitutional arrangements of postwar democracies. Yet they also coexist with forms of political constitutionalism and mixed government. It remains to explore their respective advantages and disadvantages and the tensions between them.

Political and Legal Constitutionalism Compared

An entrenched, rights-based, and justiciable constitution is said to ensure stable and accountable government, obliging legislatures and executives to operate according to the established rules and procedures, and above all prevents their sacrificing individual rights to administrative convenience, popular prejudices, or short-term gains. Given no working constitutional government has not been also a working democracy, few analysts believe constitutions can restrain a genuinely tyrannical government. Rather, the aim is to prevent democratic governments from falling below their self-professed standards of showing all equal concern and respect. So a legal constitution is seen as a corrective to—even a foundation for—a working political constitution. Yet it remains a moot point whether it performs its appointed task any more effectively or legitimately.

Democratic governments are said to be prone to overreacting to emergency situations, sacrificing civil rights to security, and pandering to either electorally important, yet unrepresentative, minorities or the populist sentiments of the majority. Insulated from such pressures, a court can be more impartial while its judgments are bound by constitutional law. However, others contend these supposed advantages turn out to be disadvantageous. Going to law offers an alternative to entering the political realm, yet access is more restricted than voting and the costs of a case as prohibitive to most ordinary citizens as founding a new party. Meanwhile, it allows those with deep pockets to fasten on to a single issue that affects their interests without the necessity of winning others to their point of view. Courts may be restricted to the law in their judgments—but what does that mean? Is the law to be found in the text of the constitution, the original intentions of those who drafted it, the objective meaning of the principles, or the common understandings of the people? Words are open to multiple meanings, so textualism hardly proves binding on judgments, while semantics seems an odd way to decide difficult moral and political issues. The intentions of the drafters are unlikely to be consistent or knowable and may well be inappropriate in contemporary conditions. Being bound by the past favors the status quo and those who are privileged by current arrangements, thereby hindering progressive reform. If the principles behind the constitution are universal and timeless, then it could be applied to any and all situations. Yet legal philosophers—no less than citizens—disagree whether such principles even exist, let alone what they might require in particular cases. Appealing to a popular consensus will not resolve that problem, for it is either unlikely or better provided by a political constitutionalism that consults popular views directly. In all these respects, judicial review risks becoming arbitrary rather than being a block on arbitrariness.

As legal constitutionalism has spread, establishing itself not just in former authoritarian regimes but also in the United Kingdom and Commonwealth countries where political constitutionalism had hitherto held sway alone, so empirical scholarship has highlighted these drawbacks. More often than not, legal constitutionalist arrangements have been introduced by hegemonic groups fearing political challenges to their position, with the record of the new regimes faring no better overall on civil rights and, from an egalitarian perspective, rather worse on social and economic rights. Whereas political constitutionalism responds to majority views for enhanced and more equal public goods, legal constitutionalism has invariably inhibited such reforms on grounds of their interfering with individual property and other rights. Nor has it upheld political constitutional arrangements particularly well—for example, blocking campaign finance limits in many jurisdictions. Of course, important exceptions exist, with the progressive rulings of the Warren Court (1953–1969) in the United States offering an apparent contrast to the free market decisions of the Lochner era (1897–1937). However, these decisions largely reflected sustained, national, majority opinion and only became effected when backed by legislative rulings and executive action. At best, legal constitutionalism proves only as good as the political constitution; at worst, it inhibits its more equitable and legitimate working.

Richard Bellamy
University College of London
London, United Kingdom

See also Constitutional Engineering; Rule of Law

Further Readings

Bellamy, R. (2007). *Political constitutionalism: A republican defence of the constitutionality of democracy*. Cambridge, UK: Cambridge University Press.

Hirschl, R. (2004). *Towards juristocracy: The origins and consequences of the new constitutionalism.* Cambridge, MA: Harvard University Press.
Vile, M. (1967). *Constitutionalism and the separation of powers*. Oxford, UK: Clarendon.

CONSTRUCTIVISM

Constructivism is a theory according to which social phenomena are constructed through interactions among humans, who interpret one another's actions and define situations based on those interpretations. Thus, constructivism offers a way of studying social phenomena, which people tend to treat as though they were objective entities. However, from the viewpoint of constructivism, what people believe to be objective entities are actually accomplished through interactions between human actors who interpret those phenomena within specific social and historical contexts.

Constructivism is not a theory composed of a series of hypotheses but a perspective that studies discourse in order to analyze phenomena. This perspective gained prominence following the publication in 1966 of Peter Berger's and Thomas Luckmann's *The Social Construction of Reality.* Since then it has become widely influential throughout the social and human sciences. For example, in anthropology and sociology, there were the debates between "essentialism" and "constructivism" concerning sex, race, and ethnicity. The debate made it clear that sex and race cannot be differentiated using only biological standards nor reduced to unchanging essences. People use these categories in practical ways, contingent on the context: Depending on the situation, a certain gender or racial category is attributed to a particular person.

Thus, it is impossible to identify sex using only an objective biological standard. There are people who experience an inconsistency between their biological sex and their subjective consciousness of the sex to which they think of themselves as belonging. In Japan, a law was passed in 2003, by which people who undergo gender reassignment surgery and who do not have any juvenile children can be categorized legally as being of their new sex by getting permission from the family court. They can marry people of their previous gender. Similar laws, some of which allow more lenient conditions for changing one's legal sex, were legislated in a number of European countries.

Race (usually understood as rooted in biology) and ethnicity (understood as cultural) also prove hard to classify. Many people are so-called mixed race, and ascribing racial categories to them is not easy. For example, the U.S. Census treats "Hispanic" and "Latino" as ethnic categories, and Hispanics or Latinos may classify themselves as belonging to the racial category of White, Two or More Races, or Some Other Race. However, in informal contexts, Latino and Hispanic may be considered to be either ethnic or racial classifications.

Ethnicity is characterized by cultural traits such as language, religion, customs, and social behavior, but standards for ascribing ethnicity also are uncertain. In the United Kingdom (UK), Chinese are sometimes considered an independent category, differentiated from the separate Asian category, while there are different ethnic groups among Arabic people. Also, as ethnicity has become the focus for many nationalist movements in the world, it becomes apparent that the concept of ethnicity itself is a historical product.

On Terminology

Readers may have encountered two terms: *constructivism* and *constructionism.* Concerning these different terms, Holstein and Gubrium, the editors of a comprehensive handbook on the study of constructivism published in 2008, point out that, although *constructivism* is the preferred term in science and technology studies and *constructionism* is more widely used in the social sciences, the two terms can be used interchangeably in most cases. Joel Best, another sociologist, notes that *constructivism* has high cultural overtones and appears to be favored by British scholars, although American sociologists seem to use the two terms interchangeably. Thus, this entry's use of *constructivism* as a generic term in the social sciences encompasses *constructionism*, the term often used in empirical research by political scientists.

Constructivist Studies of Science

No one can dispute that the roots of constructivism are in the sociology of knowledge. However, the sociology of social problems and the sociology of science were the specialties that took the lead in exploring constructivism and within which this perspective became dominant. Using ethnographical and/or anthropological studies, sociologists interested in studying science from a constructivist viewpoint described how scientists reach agreement when they discover new information.

Although these researchers focused on the interactions among scientists in the laboratory, they tended to neglect outside influences. Constructivists who study the sociology of science often do limited, microsphere ethnography in laboratories. Some of them focus narrowly on scientists' conversations, as though only discourse can determine what is considered a scientific finding. These researchers are liable to be committed to relativism. That is, some of them seem to argue that anything might be recognized as true so long as there is consensus through discourse within a community, or they seem to imply that there is no way to make sure of the certainty of assertions. When some of them expanded their research to scientists' networks outside the laboratory, they discovered much broader, structural contexts that influence the activities of scientific research.

Constructivist Studies of Social Problems

The study of social problems is the constructivist work most relevant to political science. Research in which sociologists escaped from the trap of relativism can be seen in the study of social problems. Social constructivist studies of social problems evolved from the labeling perspective on deviant behavior. Labeling, which flourished in the late 1960s and early 1970s, focused less on the deviant behavior itself than on the process through which some behavior is defined and treated as deviant. It studied how deviant behavior is socially constructed.

Positivists who study social problems search for the causes of social problems. Constructivists, however, point out that social problems are taken for granted by positivist scholars. Positivists presuppose what are social problems and treat them as though they are social conditions, rather than first studying how some phenomenon is constructed into a social problem through the interactions among social actors, agencies, groups, organizations, and institutions. A social problem is what those people—not academic researchers—*define* as a social problem. Thus, as Malcolm Spector and John Kitsuse (1977) assert, social problems are not objective conditions but the activities of individuals or groups who make claims with respect to some "putative conditions." Analysts should focus on the people making and responding to claims and counterclaims and on the interactions among claims-making groups and responding groups and institutions.

When this constructivist perspective was criticized for inevitably presupposing some characteristics of the "putative conditions," it divided into two schools: strict constructivism and contextual constructivism. Strict constructivism claims that researchers should refrain from any presuppositions about the characteristics of "putative conditions" and aims for a pure, presupposition-free constructivist position. Contextual constructivism says that it is impossible to have presupposition-free constructivism. It also says that scientists can sometimes check the "putative conditions" and find that some claims might be absurd, even though the aim of contextual constructivist research is not debunking the content of the claims. At the beginning, there seemed to be much support for strict constructivism, yet the great majority of published research adopts the stance of contextual constructivism.

Even when the same behavior is claimed to be a social problem, the people making the claims can adopt very different "frames," depending on historical and social contexts. Social problems that are constructed using frames that are easy to accept tend to become established, familiar social problems. Recent examples of such widely accepted frames include "human rights," "health," and "democracy." Consider the case of smoking. Selling cigarettes was banned in 15 states, such as Kansas, Illinois, and Minnesota, in the United States by 1909. (During World War I, people became tolerant of cigarette smoking, and these laws were later abolished.) This was accomplished by defining it as a vicious habit and an immoral behavior from the viewpoint of Protestant ethics. More recently, smoking became regulated because of its harmful effects on health. But even the health frame can be constructed in different ways: Initial claims emphasized the damage to the smoker's own health; however, more recent regulations have been justified in terms of the risks to others' health caused by passive smoking. Thus, the current construction invokes a synthesis of health and rights frames.

Once a social problem gains acceptance, it can undergo what Best (2008) calls "domain expansion." When "hate crime" was first categorized, it was defined as a crime caused by racial prejudice. However, its domain has expanded to include prejudice against sexual orientation. Similarly, child abuse initially meant physical violence inflicted on children, such as beatings, but now it has been expanded to include verbal and psychological abuse.

When the mass media focus on some phenomenon claimed to be a social problem, to attract more attention and have a stronger impact, they tend to depict an extremely serious case. Because it is referred to repeatedly, it becomes a high-profile case, and people are liable to think of this instance as though it were a typical case of that social problem. For example, claims about missing children are illustrated with cases of children kidnapped and murdered by strangers. However, the reality is that most child abductions are committed by a separated partner in the course of a family dispute, while the great majority of children reported as missing have been runaways who returned home safely and voluntarily.

Social problems are constructed by using the language and narratives of claims makers, victims, supporters, and the media and through interaction among them and their readers and audiences. First, a phenomenon should be recognized and named. In that sense, it is constructed as being linguistically different from other phenomena. Such claims serve as a kind of advertising activity, promoted by mobilizing resources to make the phenomenon recognized as a serious problem to others and to demand some sort of solution from among various alternatives. Rhetoric is crucial to stimulate people's emotions and to persuade them. Thus, victims are presented as innocent and vulnerable.

This discourse of claims can be analyzed from different approaches to constructivism. Historically sensitive constructivism inspired by Michel Foucault analyzes discourse at a macrolevel. Constructivism informed by ethnomethodology and conversation analysis studies discourse at a microlevel.

Construction of Crime Problems and Social Policy

In advanced societies, social problems and politics intersect when deviant behavior is being constructed as a criminal problem. One strategy taken to solve such problems is to "get tough" and mete out harsher punishment. This policy is most evident in the United States and the UK.

In the United States, the "broken-windows" theory has been put into practice. According to this theory, broken windows, graffiti, and similar public displays of neglected property and petty criminality tend to encourage further criminal behavior. The strategy in this theory is to regulate such minor offenses in order to prevent serious offenses. In the UK, the government has instituted laws to control antisocial behavior. Drawing graffiti, making noise to disturb neighbors, annoying pedestrians, and other similar behaviors are defined as antisocial. When people commit such behavior for the first time, they are cautioned and get an Anti-Social Behavior Order (ASBO), which leads to a civil case before a magistrate's court. If they exhibit antisocial behavior again and breach the ASBO, they can be prosecuted by an agency of the local government at a magistrate's court as a criminal case and can be imprisoned.

People are afraid of youth crime. By showing themselves to be "tough on crime," politicians gain popularity and receive more votes from their constituents. Even though politicians may know that the true dichotomy is not tough versus soft on crime, when they see the tabloid newspaper headlines and articles, which they think reflect public opinion, they hesitate to be seen as soft, being afraid to lose popularity among voters.

Media coverage can foster a strong fear of crime. Also, it encourages readers and audiences to empathize with victims. The fear of being victimized and the desire for revenge against the perpetrators rises, and people demand tougher policies against crime. Politicians not only think it their responsibility to respond to the population's desire but also try to use the situation as an election instrument. This is "penal populism," promoted by distorted public opinion inspired by the mass media's sensational crime reporting, politicians' posturing so as not to lose popularity, campaigns by interest groups including formal social control agencies, and the enterprises that run private correctional facilities.

In 2008, the International Centre for Prison Studies reported that the number of people incarcerated in U.S. prisons and jails was 2,293,157 (including pretrial detainees). The incarceration rate was 756 per 100,000 of population: the highest rate in the world, and according to U.S. Department of Justice statistics, one in nine Black males between the ages of 20 and 34 years was in jail. The International Centre reported that in England and Wales, the number of prisoners nearly doubled, from 42,000 in 1991, to more than 83,000 in 2008. The 2008 incarceration rate was 153 per 100,000 population.

Another aspect of the social construction of policy is the social construction of target populations of social policy. Research done by constructivist policy scholars shows that people who are most vulnerable tend to participate least in politics, so their interests are liable to be ignored in the designs of social policies. It is also the case that people who are deemed the target of benevolent social policies tend to be disadvantaged and therefore neglected, as they suffer from a scarcity of resources and fail to participate in politics.

Constructivism and International Relations

In the study of international politics, realism has been the dominant theory. That perspective supposes that nations unwaveringly pursue power and wealth. However, the situation of "the war of all against all" is not always a natural condition nor universally adaptive to all situations. If the leading nations adopt an attitude that assumes all other nations are enemies, the situation of the war of all against all seems to be accomplished by the reactions of other nations, which is to take the same attitude to defend themselves from stronger nations. This becomes a self-fulfilling prophecy. However, this is not an unavoidable or necessary situation. Alexander Wendt (1999) identified three types of cultural conditions of anarchy and called this type "Hobbesian." He pointed out that there are alternatives such as "Lockean," which is based on roles of rivalry, and "Kantian," which is based on friendship roles. Using the structuration theory of Anthony Giddens and symbolic interactionism in sociology, Wendt postulates that agents and structures are mutually constitutive, and he emphasizes agents' interpretive activities through which collective identities are formed.

Thus, what looks to be a general principle presupposes some conditions that are not universal. Constructivism points out what is taken for granted by ordinary citizens, diplomats, leaders, groups, organizations, agencies, and states. Thus, in the field of international relations, there cannot be objective and universal rules such as those that govern in the natural sciences.

From the viewpoint of constructivism, the actions of nations are also performed according to their accounts and interpretations of what is considered legitimate, appropriate, or authentic. It is not only material power but also ideas and norms that influence their actions. Actions of nations are not automatic reactions to the global power structure. Further, they cannot be predicted using only a rational choice theory based on the calculation of nations' material powers and interests.

Not only nations but also agencies, such as nongovernmental organizations, nonprofit organizations, human rights organizations, the United Nations (UN), the UN's Human Rights Committee, the European Union, the Council of Europe, the European Commission of Human Rights (now obsolete), and the International Criminal Court, have become more significant in international relations. It is important to take into account these groups' own definitions of the situations concerning norms, appropriateness, legitimacy, roles, and identities that affect their behavior.

Because nations interpret and define situations, they are not puppets or organs that produce the same outcome confined by the international structure. Compared with the adherents of realism, which emphasizes material power, constructivists are interested in researching international norms that affect international relations. Constructivists explore how new norms emerge, cascade, and become internalized within international relations. Although the emergence of norms in international relations is an interesting theme of constructivism, in sociology the control of human behavior by rules has been studied mainly by structural functionalists. However, in international relations, global situations are much more fluid than the social structure of a single state, and constructivists study all the facets of norms, such as their emergence, interpretation, role taking, and life cycles. This complexity is why constructivism, rather than structural functionalism, is needed to study international relations.

When a norm is applied, we can expect that interpretations of it will be very different from situation to situation. The pragmatics of norms, the rhetoric that is endorsed when norms are activated, and the discourse that can mobilize concepts such as legitimacy, authenticity, appropriateness, and conventions are interesting themes of constructivism. One of the main characteristics of constructivism is that it does not assume that the behavior of nations is objectively predetermined by material power or interest; rather, constructivism recognizes that interpretations of ideas and definitions of situations by human agencies shape international relations.

Conclusion

Constructivism has become too influential in most of the social sciences to be ignored. In some fields, such as the study of ethnicity and race, it has become almost impossible to find studies that have nothing to do with constructivism or were not inspired by it. The number of academic papers in the social sciences that contain terms relevant to social construction has increased (Best, 2008). Since the connotations of "social construction" or "socially constructed" have become so diversified in academic papers and books, it looks as if the term *constructivism* (or *constructionism*) might no longer cover all that their usages have come to signify.

Although the majority of research in political science and the social sciences continues to be within the positivist tradition, constructivism has become increasingly valued for its insights and creativity, and it will likely continue to play a growing role.

Jun Ayukawa
Kwansei Gakuin University
Nishinomiya, Japan

See also Constructivism in International Relations; Discourse Analysis; Political Philosophy; Political Theory

Further Readings

Berger, P., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. New York: Doubleday.

Best, J. (1989). *Images of issues*. New York: Aldine de Gruyter.

Best, J. (1993). But seriously folks: The limitation of the strict constructionist interpretation of social problems. In J. Holstein & G. Miller (Eds.), *Reconsidering social constructionism: Debates in social problem theory* (pp. 129–149). New York: Aldine de Gruyter.

Best, J. (2008). Historical development and defining issues of constructionist inquiry. In J. Holstein & J. Gubrium (Eds.), *Handbook of constructionist research* (pp. 41–64). New York: Guilford. Holstein, J., & Gubrium, G. (Eds.). (2008). *Handbook of constructionist research*. New York: Guilford.

Holstein, J., & Miller, G. (Eds.). (1993). *Reconsidering social constructionism: Debates in social problem theory*. New York: Aldine de Gruyter.

Roberts, J. V., Stalans, L. J., Indermaur, D., & Hough, M. (2003). *Penal populism and public opinion: Lessons from five countries*. Oxford, UK: Oxford University Press.

Spector, M., & Kitsuse, J. (1977). *Constructing social problems*. Menlo Park, CA: Cummings.

Walmsley, R. (n.d.). *World's prison population list*. London: International Centre for Prison Studies. Retrieved October 15, 2010, from http://www.kcl.ac.uk/depsta/law/research/icps/downloads/wppl-8th_41.pdf

Wendt, A. (1999). *Social theory of international politics*. Cambridge, UK: Cambridge University Press.

Constructivism in International Relations

Constructivism in international relations (IR) refers to a family of theoretical approaches that share three analytical focal points in appraising world politics: intersubjectivity, the mutual constitution of agents and structures, and the double hermeneutics (i.e., the interpretation of interpretations). Heavily inspired by sociological thinking, constructivism is not a substantive theory of international politics in the same way that neorealism is, for instance. In the IR discipline, constructivists generally seek to redress the lack of attention given to social factors in political life, which characterizes rationalist utilitarian models. Starting from the premise that world politics basically consist of social relations, constructivist scholars believe that international politics are not fundamentally different from other spheres of human activity, where practices are produced, reproduced, and contested inside a meaningful and patterned social context.

Although constructivism emerged 5 decades ago in philosophy, sociology, and anthropology, it did not reach the field of IR until the late 1980s and early 1990s. Two main evolutions favored the rise of constructivism in IR. First, the failure of dominant theories to predict or, more simply, to explain the end of the Cold War, arguably one of the most important international changes in decades, led to something of an existential crisis in the discipline. Second, the rise of the "third debate" in IR theory about epistemology shook up more orthodox understandings about how social scientists should go about the study of world politics. During the 1990s, constructivism gradually imposed itself as a via media, or middle ground, in the IR theoretical landscape, revisiting a number of disciplinary foundations with novel ideas such as Alexander Wendt's "anarchy is what states make of it" (Wendt, 1999).

One evocative way to locate constructivism in IR is to look at the notion of interest and the analytical work that it does according to different theories. Like rationalism, constructivism assumes that agents have interests and that they act on the basis of those interests. But while rationalism takes preferences to be exogenously given (i.e., the origin of interests falls outside the scope of the theory), for constructivism the question of where interests come from is front and center in the analysis. Just how do people come to want what they want? From a constructivist perspective, the crucial issue of interest formation cannot be dismissed without losing sight of a fundamental dimension of politics. If it is true that interests drive the world, then we must know where they come from and how they form. By contrast with the rationalist science of decision making, focused on how people act on the basis of preexisting preferences, then, constructivism favors the analysis of sense making—that is, how people define and construe their interests. Instead of being magically read off a material structure, interests are contextually defined (Focal Point 1: intersubjectivity); they are not individually defined by atomized individuals but generated in and through social relations (Focal Point 2: mutual constitution of structure and agency); and they need to be interpreted, at the levels of action and observation (Focal Point 3: double hermeneutics).

The following portions of this entry introduce constructivism in IR in two different ways. The first section reconstructs the three developmental stages that the family of theories has gone through over the past 20 years. The second section discusses and illustrates the three conceptual building blocks that constructivism is premised on. The conclusion assesses whether constructivism, one generation after its rise, has lived up to its promises to open new thinking space in the IR discipline.

The Three Moments of Constructivism in International Relations

The word *constructivism* appeared on the IR radar screen for the first time in Nicholas Onuf's 1989 book *World of Our Making*. In the following decade, the theoretical label quickly gained in prominence, experiencing its heyday at the turn of the millennium with the publication of Wendt's seminal *Social Theory of International Politics* (1999). Now a vibrant approach in the IR theoretical landscape, constructivism has shifted gears lately, taking a break from metatheory to move into questions of applied research and methodology. Twenty years after its rise, constructivism is not an exotic approach any more in IR, as the disciplinary *aggiornamento* that accompanied its rise came to pass. To be sure, as is always the case with historiography, this narrative simplifies and rationalizes post hoc a much more complex and contested disciplinary trajectory. Its heuristic purpose is limited to understanding better what constructivism does differently from other IR approaches, by grasping the evolutionary stages that the approach went through over the past 20 years: an epistemological moment, an ontological moment, and a methodological moment.

An Epistemological Moment

The "third debate" in IR, which emerged in reaction to the neorealist orthodoxy of the early 1980s, was the necessary condition for constructivism to emerge. Influenced by continental philosophy, a number of critical scholars started to question the rampant positivism of the IR discipline. Building on the insights of the linguistic turn (a development in 20th-century thought that focused on the relationship between philosophy and language), they interrogated the scientific ambition by emphasizing the performativity of language and, by implication, of theory and practice. In replacement for the naturalist orthodoxy, critics promoted a postpositivist epistemology premised on reflexivity and the plurality of viewpoints and methods. More than 20 years later, it is fair to say that the third debate in IR was inconclusive, with both sides retreating to their respective trenches with many bruises but little advancement of their positions.

In this context of disciplinary upheaval, constructivism emerged as a form of epistemological third way that had, for a time at least, much to offer to both positivists and postpositivists. For the former, constructivism facilitated the cooptation of certain novel insights into the IR theoretical mainstream—for instance, about the role of ideational variables. For the latter, constructivism's rising profile in IR appeared to grant some legitimacy to a large part of the postpositivist epistemological critique. In the end, however, the middle ground proved a very uneasy position to occupy, and fierce debates over "explaining versus understanding" continued to rage among IR scholars. To this day, there still exists a fracture between (scientific) realist constructivists, whose epistemology can accommodate positivism, and interpretive constructivists, who are philosophically closer to postpositivism. Though inconclusive, the epistemological moment of constructivism came to pass during the 1990s as key proponents of the approach moved the spotlight onto ontological matters.

An Ontological Moment

Born out of an epistemological controversy, constructivism imposed itself in IR largely by shifting attention away from the third debate and toward ontology. Building on scientific realism, several constructivists argued that the "stuff" of reality that is out there should take analytical precedence over the way that we get to know it. Ontology, in other words, matters more than epistemology. While this position was crucial in establishing IR constructivism as a "legitimate" approach in the eyes of the disciplinary mainstream, among constructivists it has been hotly contested ever since Wendt and others put it forward. The majority view nowadays rather portrays ontology and epistemology as two sides of the same coin: After all, as Stefano Guzzini notes, constructivism is about the social construction of knowledge *and* the construction of social reality.

Constructivism brought two key ontological innovations to IR, which are discussed at greater length below. First, building on the sociology of knowledge, constructivism emphasizes the social construction of reality: Ideas, meanings, and other forms of collectively held beliefs matter in the explanation of social and political life. In IR, that argument is a direct rejoinder to neorealism, by which ideas are nothing but epiphenomena of deeper material forces, as well as to cognitivism, which tends to reduce beliefs to ideas held by individuals. Second, taking inspiration from developments in sociological theory, constructivism puts the mutual constitution of agents and structures at the center of analysis. Against both individualism and holism, constructivism argues that actors and their environments determine one another in a recurring pattern of coevolution. Building on these two innovative insights, IR constructivists have sought to rejuvenate the ontological foundations of the IR discipline by putting forward a number of new concepts and by revisiting old ones such as norms, identities, rules, communicative action, culture, and communities. Pointing to the many socially constructed realities that make up world politics, constructivists throughout the 1990s legitimized their frameworks by "discovering" novel ontological entities—the most famous one being Wendt's 1999 reinterpretation of anarchy as a cultural and historically contingent structure of interaction.

A Methodological Moment

Despite its success, the ontological revolution that constructivism brought about in IR did not go without problems. First, it had been made possible by brushing under the carpet a number of important epistemological controversies. To take an obvious one, if reality is socially constructed, what are the implications for the development of social-scientific knowledge? And second, the success of constructivism sparked a number of defensive reactions from the disciplinary mainstream. By the beginning of the 21st century, a metatheoretical fatigue had set in as calls for constructivists to do "real research" and study actual problems in world politics multiplied.

Since the turn of the millennium, IR constructivism has focused largely on issues of methodology, operationalization, and empirics. How should constructivism proceed with the study of world politics given its ontological innovations and epistemological diversity? In the past decade, the number of in-depth empirical analyses of world political phenomena has exploded. That evolution is paralleled by an increasingly loud debate about methods. What is discourse analysis and how should it be practiced? Is quantification amenable to the constructivist style of reasoning? What type of fieldwork data can best support a constructivist argument? How can we adjudicate between explanations—constructivist or otherwise? Can constructivist findings travel from one case to another? Often grounded in the empirical analysis of international politics, these very important methodological issues have received increasing scrutiny over the past few years, allowing constructivism to enter a constructive and grounded dialogue with the rest of the IR discipline.

The Three Building Blocks of Constructivism in International Relations

What do constructivists do, as they research world politics, that makes them distinct from scholars who favor other theoretical approaches in IR? There are obviously many ways to answer this question. In the following paragraphs, constructivism is characterized by the three analytical focal points that its advocates typically put to work in their analyses: (1) intersubjectivity, (2) the mutual constitution of structure and agency, and (3) the double hermeneutics.

Intersubjectivity

Intersubjectivity refers to the condition of meanings that do not depend on a particular point of view to exist. In an intuitive sense, ideas belong to individuals and are located between people's earlobes. Taking this view, a number of IR specialists have emphasized the role of perceptions in shaping world politics. But constructivism stresses a different (though related) kind of meaning—not subjective but *inter*subjective. Contrary to ideas held by individuals (e.g., a cognitive bias), intersubjective meanings define reality as something independent of our volition. The classic example, used by John Searle (1995), is that of money: In a banking system, certain bits of paper engraved with specific markings are consensually taken to be worth 20 dollars. This socially constructed meaning, which is attached to bits of paper by a collectivity, is very real in its effects (e.g., in buying groceries), regardless of one's personal misgivings about capitalism and despite the fact that for Martians paper money would probably be worthless. Conventions—just like rules, norms, identities, cultures, or languages—are intersubjective structures, that is, objectified sets of meanings that order social configurations irrespective of what the specific agents that take part in them believe. Such is also the case, for instance, in a football game: Whatever specific strategies individual players may have, the rules of the game will define their interaction along ideational constraints that do not depend on any player's individual point of view to exist. Social facts exert effects on politics by virtue of a critical mass of relevant agents taking their reality for granted. As such, they make social action possible by creating elements of a common world.

Building on this insight, IR constructivists have focused attention on a wide variety of social facts and artifacts in international politics. International threats, for instance, are not self-evident physical facts but socially constructed realities. Similarly, norms, identities, institutions, rules, cultures, practices, languages, ideologies, and narratives are all forms of intersubjective meanings that shape world politics and guide action. States' national identities constitute foreign policies; international cultures of anarchy drive interstate relations; world structures of rules and norms determine appropriate behavior on the international stage. As John Ruggie argues, for instance, it is not simply hegemony or preponderance in material power that explains the content of our contemporary world order but the social purposes and norms that intersubjectively define the identity of the American hegemon. All in all, taking intersubjectivity seriously implies not only that meanings matter but also that they matter as structural forces in world politics.

The Mutual Constitution of Structure and Agency

The mutual constitution of structure and agency refers to the dynamic processes through which actors and contexts coevolve and codetermine one another. Inspired by Anthony Giddens's structuration theory and Pierre Bourdieu's theory of practice, the notion stresses that the social world comprises neither freestanding individuals nor free-floating structures. Instead, structures are made possible by agents' practices, which are themselves constituted by their context. According to Wendt, the idea can be simplified with two "truisms" about social life:

> 1) Human beings and their organizations are purposeful actors whose actions help reproduce or transform the society in which they live; and 2) society is made up of social relationships, which structure the interactions between these purposeful actors. Taken together these truisms suggest that human agents and social structures are, in one way or another, theoretically interdependent or mutually implicating entities. (Wendt, 1987, pp. 337–338)

In other words, structures emerge out of agents' practices and vice versa; neither is reducible to the other. Note that this view entails a processual ontology by which agents and structures should not be treated as reified or static substances but as evolving sets of relations and interplays of practices. As stabilized as certain patterns may be, any practice either reinforces or undermines the existing order of things. Agency, the power to deviate from structure, introduces contestation as well as contingency.

To grasp what the mutual constitution of structures and agents changes for the study of world politics, it is convenient to compare constructivism with rationalist frameworks in IR. The latter ontology, based on utilitarianism, envisions an atomistic set of self-constituted and self-regarding units whose identity is given and fixed. In this framework, structural effects are limited to regulating behaviors because units preexist structures. In the constructivist view, structures also affect the properties of agents, constituting their identities and interests in the first place. This makes for a much deeper structural effect: Agents would not be what they are but for structural constitution. A classic example of this constitutive effect can be found, as noted by Searle (1995), in the practice of playing chess. The rules of the game do not simply regulate the movement of the pieces on the board; they make the game possible in the first place. Without these constitutive rules, one cannot play chess; at the same time, without the application of these rules in and through practice, there cannot be an intersubjective structure called the game of chess. Similarly, the norm of sovereignty makes the current international society possible. Given the importance of constitutive theorization, the next challenge for constructivists is methodological: If intersubjective structures and meaningful practices make world politics possible, how can we account for them?

The Double Hermeneutics

Double hermeneutics refers to a methodology centered on the interpretation of interpretations. It is central to the constructivist approach in IR because world politics presents itself as a reality that is already interpreted by its actors. Since human beings are meaning makers, their interpretations are the primary object of study for constructivists. People are of course not like rocks: They attach meanings to the different parts of their environment and act on the basis of those meanings. The first interpretive moment, then, happens at the level of action. The second interpretive moment takes place at the level of observation, as social scientists seek to establish the meanings of the practices that they observe. In summary, as Giddens notes, the double hermeneutics captures the core of the social-scientific enterprise from a constructivist perspective, which is to develop scientific interpretations of lay interpretations.

In the practice of research, the notion is also reminiscent of the hermeneutic circle, according to which making sense of something involves relating the parts in terms of the whole and vice versa. Hans Georg Gadamer famously compared interpretation to a "fusion of horizons," drawing an analogy with the practice of attaining fluency in a second language to better translate from that language into one's mother tongue (Mark Neufeld, 1993). Because of interpretation, scholars cannot construe their task as simply improving the match between the world and our knowledge about it (i.e., the correspondence theory of truth). As Patrick Jackson has observed, taking the double hermeneutics seriously entails that validity does not simply stem from matching the world with words, because words partake in constituting the world. After all, most of the time, it is language that makes intersubjectivity possible by "detaching" the meaning from immediate expressions of subjectivity and representing them as "facts" external to agents. If discourse is constitutive of reality, then the social-scientific discourse, with its interpretation from afar of lay interpretations, also bears tremendous normative consequences. This is perhaps why the majority of IR constructivists exhibit a strong disposition toward reflexivity, trying to make strange what seems obvious, problematize the taken-for-granted, and denaturalize alleged universal truths. In this view, the task of IR scholars is not to define which international threats are "really real" and which ones are not, for instance; instead, they should document the political technologies that make certain threats look "really real" to various publics and should analyze the consequences that ensue for world politics.

Conclusion

This entry has argued that constructivism in IR is a broad family of theories whose distinctiveness stems from their common use of three crucial notions in analyzing world politics: intersubjectivity, the mutual constitution of agents and structures, and the double hermeneutics. In historiographical terms, constructivism has attained a degree of disciplinary legitimacy in IR by going through three stages of development: (1) an epistemological moment, which remains largely inconclusive to this day; (2) an ontological moment, which has been successfully incorporated by large sections of the discipline; and (3) a methodological moment, which is currently underway. One generation later, the "constructivist turn" has now been taken in IR, and while the approach remains far from dominant in the discipline, it cannot be relegated to the margins anymore. Constructivism is evidently here to stay. So, as an interim assessment 20 years on, has the constructivist project lived up to its promises in IR?

To be sure, the fact that constructivism has not provided a unified and applied theory of international politics in the way that Kenneth Waltz's neorealism has cannot be construed as a failure. In fact, the pluralism that characterizes the approach today speaks of its vibrancy and its capacity to accommodate theoretical debates. Since conversations among scholars hold the best hopes of pragmatic progress, the lack of orthodoxy should be welcome. In addition, the absence of one overarching and exclusive theoretical narrative is quite coherent with constructivist premises, including its critical disposition toward the imposition of the order of things. So far, so good.

But that may well be where the good news ends, because one outcome of the constructivist turn in IR that is more problematic, one could argue, is the further polarization of a discipline that was already suffering from a particularly debilitating degree of fragmentation. To caricature a bit, constructivism has become so mainstream in most parts of Europe that it is now risking redundancy: World politics is socially constructed—so what? By contrast, in many North American universities, constructivism has become eminently suspect and the target of gatekeeping strategies. This polarization is obviously very damaging for interparadigmatic dialogue. While the responsibility for this regrettable state of affairs does not fall on constructivism alone, it is hard not to take some share of the blame for the metatheoretical fatigue that now plagues the IR discipline. In this context, there may be a danger, much to constructivism's detriment, that the window of opportunity for theoretical rejuvenation in IR that opened with the end of the Cold War is now shutting down.

Vincent Pouliot
McGill University
Montréal, Québec, Canada

See also Anarchy; Constructivism; Critical Theory in International Relations; Hermeneutics; International Relations, Theory; Logic of Appropriateness

Further Readings

Adler, E. (2005). *Communitarian international relations: The epistemic foundations of international relations*. New York: Routledge.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Cambridge, UK: Polity.

Guzzini, S. (2000). A reconstruction of constructivism in international relations. *European Journal of International Relations,* 6(2), 147–182.

Jackson, P. T. (2008). Foregrounding ontology: Dualism, monism, and IR theory. *Review of International Studies, 34*(2), 129–153.
Klotz, A., & Lynch, C. (2007). *Strategies for research in constructivist international relations*. Armonk, NY: M. E. Sharpe.
Neufeld, M. (1993). Interpretation and the "science" of international relations. *Review of International Studies, 19*(1), 39–61.
Onuf, N. G. (1989). *World of our making: Rules and rule in social theory and international relations.* Columbia: University of South Carolina Press.
Pouliot, V. (2007). "Sobjectivism": Toward a constructivist methodology. *International Studies Quarterly, 51*(2), 359–384.
Ruggie, J. G. (1998). *Constructing the world polity: Essays on international institutionalization*. New York: Routledge.
Searle, J. R. (1995). *The construction of social reality.* New York: Free Press.
Wendt, A. (1987). The agent–structure problem in international relations theory. *International Organization, 41*(3), 335–370.
Wendt, A. (1999). *Social theory of international politics.* New York: Cambridge University Press.

CONTAINMENT

The concept of containment was presented by the American diplomat George F. Kennan in a long telegram to the U.S. Department of State on February 22, 1946, and in an article published in the July 1947 issue of *Foreign Affairs*, signed "X." The concept was intended to influence the U.S. policy response in the specific strategic context of the Cold War, but with the collapse of the Soviet Union in 1991, the end of a bipolar world raised new questions regarding the relevance of containment—especially as the George W. Bush administration began to target new kinds of enemies.

The Cold War and the Origins of the Concept

Containment was adopted by President Harry S. Truman's administration (1945–1953), both as a doctrine and as a rationale for external action, and was carried on by Truman's successors, especially Lyndon Johnson (1963–1969). The main objective of this new doctrine was to use military, economic, and diplomatic means to oppose what Kennan depicted as the Soviet Union's "hegemonic" strategy: "The main element of any United States policy toward the Soviet Union must be that of a long-term, patient but firm and vigilant containment of Russian expansive tendencies" (X [Kennan], 1947, Part 2). The United States pursued a host of policies in the name of keeping an essential commitment to prevent the spread of the Soviet influence throughout the world. A large proportion of these actions were carried out in Europe and Asia. Some examples include support given to Greece to fight against the "Communist subversion" (1947); the launching of the Marshall Plan (a program of direct economic aid to Europe); the promise to support "free peoples who are resisting attempted subjugation by armed minorities or by outside pressures," as stated by President Truman in a speech on March 12, 1947; the construction of the North Atlantic Treaty Organization (NATO) (1949); and the Korean and Vietnam Wars.

Although declining in the late 1960s (during the "détente" promoted by U.S. President Richard Nixon and the U.S. Secretary of State Henry Kissinger, then during the easing of tensions with the USSR and communist China), and challenged by the more radical concept of "rollback" (most notably under President Eisenhower's secretary of state John Foster Dulles, who called for the "liberation" of Eastern Europe), the policy of containment continued to mark the American foreign policy landscape until the end of the Cold War, as President Ronald Reagan (1981–1989) sent military aid to anticommunist movements in Afghanistan, Angola, Cambodia, and Nicaragua; deployed the Pershing II missiles in Europe; and promoted the Strategic Defense Initiative, which would use ground- and space-based systems to protect the United States against nuclear missile attack.

Both external and internal factors influenced the relative power of containment as a doctrine. The Truman administration had the advantage of a favorable economic context. Soviet Premier Joseph Stalin's death in 1953 allowed U.S. President Dwight D. Eisenhower to be less influenced by balance of power relations. Instead, he focused more on cutting government spending and reducing economic assistance plans and gave less priority to expensive conventional military deployments. It would have thus been reasonable to deduce that the end of the Cold War (1989–1991) would have marked the official end of the United States' reliance on containment policy. Once America's 40-year enemy disappeared, the expectation was that a "brave new world" would pay "dividends of peace" and render obsolete an expensive containment policy. But 20 years after the fall of the Berlin Wall, there are more than 700 operable U.S. military bases, and approximately 370,000 American troops remain deployed in more than 150 countries. Nevertheless, fears continue to be expressed about the need to circumscribe emerging peer competitors, new global enemies, and new threats.

Post-Bipolar Containments?

If the rhetoric of containment has officially disappeared with the demise of the Soviet Union, its spirit can be found in the targeting of new "threats" by the United States, especially in three main categories: rogue states, peer competitors, and ambiguous entities.

Rogue States

Iraq was the subject of a containment policy between the 1991 Gulf War (following its invasion of Kuwait) and the attacks on the United States in September 2001. Severe sanctions were applied, United Nations weapons inspections were imposed, U.S. troops and air patrols in the Persian Gulf were deployed, and the economic weakening and political isolation of Iraq was put into place. The uniqueness of the Iraqi example stems from the fact that it led to a radical change of course when the administration of George W. Bush opted for a policy of regime change via military action in 2003 and toppled Saddam Hussein. This raised questions about whether a similar fate awaited Iran and North Korea, the two other states (besides Iraq) characterized in Bush's 2002 State of the Union address as the "axis of evil" and potential nuclear threats.

Peer Competitors

A new containment of China was hinted at (but officially denied) by U.S. Secretary of State Condoleezza Rice as she paid tribute to Kennan (March 2005 in Tokyo) following his death. The context was the elaboration of a new logic of maintaining alliances, along with a mentioning of the need for China to play "a positive role." In 2009, the United States had military bases in South Korea, Japan, and Afghanistan, much as it provided military equipment to South Korea, Japan, India, and Taiwan (whose security is guaranteed by Washington). Even after the election of Barack Obama to the presidency in November 2008, declarations by the new Secretary of State, Hillary Clinton, evoked the need to diminish China's global economic influence. Russia, after invading neighboring Georgia in August 2008, was also considered to be somewhat of a threat requiring containment: Moscow felt directly targeted by the installation of a U.S. missile defense shield in Poland and the Czech Republic (even though Washington insisted its official purpose was to counter "rogue states" such as Iran).

Ambiguous Entities

A relevant innovation following the terrorist attacks on September 11, 2001, was the effort by the United States to drag its allies into a third type of containment against "terror." Identified by Washington as the new global security priority since 2001, the "war on terror" became a massive endeavor to counter political or religious groups and movements—even social, economic, or cultural practices—with potential links to terrorism. With efforts concentrated mostly in the Middle East and the Muslim world (in Iraq, then Afghanistan; in the Gulf but also in Central Asia, Southeast Asia, and Africa), the United States has been deploying troops and military bases, fighting against social and financial networks, and working to change political regimes and to control transnational flows.

The revelation found in this new type of containment lies in the nature of the enemy targeted. With neither a clear political nucleus nor a defined territory (even if linked to the Muslim world in American rhetoric), the "terrorist" entity hardly conforms to Kennan's initial containment theory, which was proposed to counter the Komintern and which included cutting diplomatic relations and all exchanges with the state in question. Still, Washington's insistence on organizing the eradication of both "terror" and its state and nonstate sponsors recalls the initial containment doctrine and raises the question of whether the concept of containment is enduring, relevant only to a specific time, or altogether obsolete.

Containment Versus Engagement

Was containment policy efficient? And can it be effective today? In his 1947 article, Kennan advocated "the adroit and vigilant application of counterforce at a series of constantly shifting geographical and political points" by the United States. He believed that a policy of containment would add to internal pressures on the Kremlin that would ultimately lead to "the breakup or the gradual mellowing of Soviet power." Because the Soviet Union collapsed in 1991, the containment policy would seem to have proven successful. Still, the approach was controversial from the beginning and remains so.

In the 1940s, the lauded columnist Walter Lippmann put forth a harsh criticism against the "X" article for failing to differentiate between vital and peripheral American interests. Moreover, the debate persists on whether the end of the Cold War was the direct result of U.S. foreign policy strategies or whether it owes more to complex interaction between domestic and global economic and social factors.

The containment strategy now appears to be outdated for several reasons. First, containing one specific enemy in a globalized world is an illusory goal—the possibilities of circumventing any attempts at isolation are now virtually infinite. Second, containment has been used almost exclusively by the United States and its allies, and it led to a counterproductive result: States considered to be the main enemies of the United States may gain solidarity from other "protest diplomacies" and the support of public opinion (mostly in the South). Third, political and economic isolation has proved useless, and even dangerous, in an interdependent world, notably when the targeted country or entity has already acquired enough potential to cause havoc.

In contrast, dialogue and diplomatic commitment have emerged as more serious means of dealing with difficult interlocutors. Containment of China, Iran, North Korea, Venezuela—even Hamas or Hezbollah—would indeed be hazardous from material, social, political, and military perspectives. Finally, the post-bipolar structure of the international system no longer allows for such an approach; for containment was forged in a different international system with one permanent peer competitor in mind and not conceived for fast-changing and asymmetrical targets in a turbulent world.

The formal temptation of a containment policy is still present and possible. Its reemergence is possible in the context of a new global competition between the United States and China. However, in material, political, and social terms, containment is no longer tenable. Classical in its conception (a struggle for survival against an enemy in a balance-of-power competition), ambitious in its implementation, and influential among U.S. allies, the concept of containment is a remnant of the Cold War whose translation to modern times is a troubled one.

Frédéric Charillon
Université d'Auvergne–Clermont Ferrand 1
Paris, France

See also Bipolarity and Multipolarity; Cold War; Foreign Policy Analysis; Security and Defense Policy; War and Peace

Further Readings

Anderson, S. (2005). *Condemned to repeat it: Lessons of history and the making of U.S. Cold War containment policy*. New York: Lexington Books.

Gaddis, J. L. (2005). *Strategies of containment: A critical appraisal of American national security policy during the Cold War*. Oxford, UK: Oxford University Press.

Kennan, G. (1946). *The long telegram*. Retrieved October 27, 2010, from http://www.gwu.edu/~nsarchiv/coldwar/documents/episode-1/kennan.htm

Litwak, R. (2000). *Rogue states and U.S. foreign policy: Containment after the Cold War*. Washington, DC: Woodrow Wilson Center Press.

Shapiro, I. (2008). *Containment: Rebuilding a strategy against global terror*. Princeton, NJ: Princeton University Press.

X (George Kennan). (1947). The sources of soviet conduct. *Foreign Affairs, 25*(4), 566–582.

Contingency Theory

Universalist theories of organization say that the highest-performing organizations have maximum levels on the variables that compose their structure, such as maximum decentralization of decision making. In contrast, contingency theory says that the highest-performing organizations have levels on their structural variables that fit the levels on their contingency variables—that is, the level of decentralization that fits the level of organizational size. Thus, the optimal structural level is seldom the maximum level, and it is not universal across all organizations. Rather, the level of the structure that causes highest performance is contingent on the contingency variables (or factors). In other words, the contingency variable moderates the effect of structure on performance. Performance is high only if structure fits the contingency variable. The more that structure misfits the contingency variable, the lower is performance in consequence. This entry discusses the intellectual backgrounds of contingency theory and its application to the public sector.

Key Variables for Performance

Specialization, formalization, decentralization, and divisionalization are some of the variables of organizational structure that have contingent effects on performance. Some of the contingency variables of organizational structure are size, diversification, and task uncertainty. For instance, size is a contingency of decentralization, in that the larger the organization is (i.e., the more members it has), the more complex are the issues in managing it; therefore, top management cannot make all the decisions, and some must be delegated down the managerial hierarchy. This reduces the number of levels in the hierarchy through which information must pass and brings decision making closer to the bottom level, which interacts with customers or clients and which produces products.

As another instance, diversification is a contingency of divisionalization. Undiversified (i.e., single product or service) organizations can be effectively organized as structures in which the main subunits reporting to the chief executive officer (CEO) are functions. In contrast, diversified (i.e., multiple product or service) organizations can be effectively organized as structures in which the main subunits reporting to the CEO are divisions, each containing its own operating functions (e.g., sales and production), so that each is an autonomous organization. In this way, diversification is also a contingency of decentralization.

Hence, both size and diversification are contingencies of decentralization. Thus, some structural variables have more than one contingency variable they need to fit to have high performance.

Task Uncertainty

Task uncertainty is a contingency of formalization, specialization, and decentralization. Where the task has high certainty, rules (formalization) can be formulated that provide guidance for effective decisions. The preprogramming of decisions also allows jobs to be narrowly defined and hence highly specialized. In contrast, where tasks are highly uncertain, it is not possible to formulate rules that provide guides for effective decisions, and jobs cannot be narrowly defined; hence formalization and specialization need to be low. Instead, more highly qualified and skilled employees need to be empowered to make some of the decisions, so that decentralization is higher. The higher the task uncertainty, the lower the formalization and specialization and the higher the decentralization need to be in order to fit the task to the uncertainty contingency.

Uncertainty enters tasks in a major way when organizations seek to innovate, such as by creating new products or services or by using new technologies for producing its products or services. Organizations repeatedly producing the same products and services are fitted with a mechanistic structure, which is high on formalization, specialization, and centralization. Organizations innovating their products and services are fitted with an organic structure, which is low on formalization, specialization, and centralization. Innovating organizations also need more interaction between their functional departments. This requires lateral coordination mechanisms such as cross-functional project teams headed by integrators (project leaders independent of the functions from which the team members are drawn). In contrast, noninnovating organizations can adequately manage the lesser interaction between their functional departments by hierarchy and planning.

Fits and Misfits

Some form of contingency theory conceives of the fits as configurations or gestalts—that is, as bundles of structural variables where the levels of the structural variable are discrete. For instance, an organization that is small is fitted by a simple structure that is low on formalization, while a large organization is fitted by being high on formalization. These fitting levels of formalization differ considerably on any scale quantifying formalization. The intermediary levels of formalization between low and high are all misfits that cause low performance. Hence, an organization will tend not to remain at intermediary levels of formalization. Therefore, organizations make quantum leaps from low to high formalization but do so only infrequently. This is also the case for other structural variables such as specialization and decentralization, so that organizations are composed of bundles of such structural variables—the configurations or gestalts. Empirically, analysts quantify misfit by calculating the exact score on each structural variable that is the ideal for that configuration and then measure the distance of each organization from that ideal.

In contrast, one form of contingency theory conceives of the fits in a Cartesian manner. The contingency and structural variables are both continua and so are their fits. There is a line of fitting points such that each level of the contingency variable has a level of the structural variable that is its fit. For instance, size and formalization are both continuous variables, being the contingency and structural variables, respectively. Small size is fitted by low formalization and high size by high formalization (as in configuration theory), but also medium size is fitted by medium formalization, and all the intermediary levels of size are fitted by intermediary structural levels (unlike in configuration theory). Thus, there are many fits in Cartesian contingency theory, forming the fit line. Therefore, a growing organization can readily move from one level of structure that fits it, to the next level. Thus, the fit line is a set of stepping stones that allow organizations to change incrementally and gradually, thereby possibly growing from small and unformalized to medium sized and medium formalized or, eventually, to large and highly formalized. Empirically, analysts quantify misfit by locating the fit line, from theory or data, and then measuring the distance of each organization from that fit line.

The internal distribution of power between departments at the same hierarchical level addresses an aspect of the politics within organizations. The strategic contingencies theory of intraorganizational power says that the department that successfully deals with the key challenge to the organization will become the most powerful department. However, departments become powerful only if they are nonsubstitutable, so that the organization is dependent on them. For example, in brewing firms, where marketing was severely constrained by government regulations about pricing and the like, financial success came to those whose manufacturing departments produced at low cost. Consequently, these manufacturing departments received more of the budget in their firms and also controlled decisions pertaining to their boundaries with other departments—for example, the packaging department. In other industries, powerful departments such as these have been shown to supply the next CEO.

Contingency theory is functionalist in that it explains the structures that are adopted as being those that fit the contingency and so produce higher performance. It is also positivist in that it features quantitative evidence and general models, in which factors of the situation, such as the environment or size, determine the structure, so that there is a contingency imperative. Some critics assert, in contradistinction, that managers exercise free will, unconstrained by contingency variables. Others, more moderately, hold that organizations whose level of the contingency variable have changed and so are misfitted by the existing structure nevertheless retain it until there is a crisis of poor organizational performance, when they move into fit. Again, there is an argument that organizations can regain fit by adjusting the contingency variable to fit the structure. However, empirically, organizations usually change their structure to fit the contingency, so there is a contingency imperative.

Overall, there has been much empirical research supporting contingency theory, across organizations of many types, industries, and nationalities. However, some issues, such as configuration versus Cartesiansim, or contingency imperative versus free choice, remain in contention, with evidence for and against.

Traditionally in contingency theory, fit is seen as *iso-performance*, meaning that the high level of performance produced by the fit of, for instance, low formalization to small size is the same (hence "iso-") as the high level of performance produced by the fit of high formalization to large size. But this would provide no incentive for firms to grow and increase their level of formalization. In contrast, *hetero-performance* says that the level of performance produced by the fit of, for instance, high formalization to large size is greater than the performance produced by the fit of low formalization to small size. Thus, organizations have an incentive to grow and increase their level of formalization. Whereas in the traditional iso-performance theory, the contingency variable plays the role of only being a moderator of the effect of structure on performance, hetero-performance theory sees the contingency variable as also making a direct, positive contribution to performance. While hetero-performance has logical appeal, it is at the present time only a theoretical conjecture that has yet to be proven. Moreover, some contingency variables are beyond managerial control, such as uncertainty in the environment, so that the environmental contingency can increase its level without any action by managers in an organization or any incentive on them to do so. Thus, while some contingency effects are probably hetero-performance, some are iso-performance.

Traditionally again, contingency theory holds that underfit has the same effect as overfit, in that too little structure (e.g., formalization) for a given level of the contingency (e.g., size) decreases performance by the same amount as too much structure. However, it has been argued that underfit produces worse performance in that the organization is unable to attain its goals, whereas in overfit the organization can attain its goals but just incurs the unnecessary costs of an overelaborate structure.

Research is ongoing in contingency theory, giving rise to new theoretical developments and empirical findings. It represents a continuing pattern of development in the study of organizations in both the public and private sectors.

Lex Donaldson
University of New South Wales
Sydney, New South Wales, Australia

See also Decentralization; Functionalism; Organization Theory; Positivism

Further Readings

Child, J. (1975). Managerial and organizational factors associated with company performance, part 2: A contingency analysis. *Journal of Management Studies, 12,* 12–27.

Donaldson, L. (2001). *The contingency theory of organizations.* Thousand Oaks, CA: Sage.

Keller, R. T. (1994). Technology-information processing fit and the performance of R&D project groups: A test of contingency theory. *Academy of Management Journal, 37,* 167–179.

Klaas, P., & Donaldson, L. (2009). Underfits versus overfits in the contingency theory of organizational design: Asymmetric effects of misfits on performance. In A. Bollingtoft, D. D. Hakonsson, J. F. Nielsen, C. C. Snow, & J. Ulhoi (Eds.), *New approaches to organization design* (pp. 147–168). Dordrecht, Netherlands: Springer.

Lawrence, P. R., & Lorsch, J. W. (1967). *Organization and environment: Managing differentiation and integration.* Boston: Harvard University, Graduate School of Business Administration, Division of Research.

Contract Theory

The expression *contract theory* refers to a large variety of conceptions in social sciences. In the vast literature dedicated to the topic in economics, contract theory consists of the study of microeconomic relations between agents within a market. In law, contract theory is equally important and pertains to the analysis of contract law. But the most dominant usage of the expression *contract theory* is political and this is also the sense that will apply in this entry.

Within the realm of political theory, contract theory refers to a set of philosophical ideas, mostly developed since the 17th century, the goal of which is to offer a rational understanding of the creation of society and political powers. It also induces a normative reflection on the legitimacy of powers. A subset of preoccupations derives from this broad consideration and regards as crucial for political theory topics such as the place of religion in a collectivity, the question of the sovereignty, the respective role of the people and the prince, and basic rights and liberties. Beyond the borders of political thought, contractarian theoreticians have often addressed issues relevant to psychology, anthropology, law, and economics.

The core of contract theory, in this political sense, is the fundamental intuition that all social dealings are based on an initial agreement between individuals to belong to the same society and to obey certain designated authorities. Hence the idea of a "contract" at the root of social existence—often known as the "social contract" after Jean-Jacques Rousseau's famous book, *On the Social Contract*, published in 1762. According to Rousseau (1712–1778), the ties that bind individuals within a society are the outcome of a contract between them that was initiated at the beginning of this very society. Regarding the contract itself, it is a deliberate decision of each individual who sees the value of being engaged in society where rules limit liberty rather than being left alone in total freedom. This is why contract theories are all based on the idea of a reasoning individual able to recognize the greater benefits of society.

Whether this contract is regarded by authors as actually having been sealed in the course of history or whether it is for them only a fruitful hypothesis, the basic principle of contract theory is to regard society as the result of such a pact between individuals. Further, contract theory purports structural consequences such as individual rights and liberty. Indeed, if individuals had no rights, they could not logically engage in a contract (they would literally not have the right to do so). And if the contract they are engaged in did not guarantee some rights or liberty—in other terms, if social existence meant the total alienation of individuals—it would be impossible to understand why individuals would have chosen to submit to such an unfair contract and on what basis it should subsist.

Contract theory is quintessentially modern. It reflects the decisive evolution of Western political thought from the 16th century on toward an "individualistic" understanding of society and politics for which any collectivity is regarded as the distant eventual result of individual decisions. As a matter of fact, contract theory is not the only modern political conception able to accommodate individualism. But it has been one of the most prominent: If the rise of revolutionary doctrines challenged the contractarian approach in the 19th and 20th centuries, it never totally superseded it. It then gained renewed attention after the works of the philosopher John Rawls (1921–2002) in the 1970s, which insisted on the necessity of having common principles of justice agreed on in a well-ordered society. His influential *Theory of Justice*, published in 1971, claims the heritage of the tradition of the social contract and has had a considerable influence, which still continues to feed contemporary political thought.

Normative Diversity and the Intellectual Unity of Contract Theory

The term *contract theory* did not come into widespread use until the 20th century, and the expression should be considered as a flexible label by which to categorize a tradition of thinkers among whom influences are easily spotted, even across centuries, but who are also very different. Their very understanding of what the "contract" at the origin of society is, the depiction of its nature, and the role it plays in their conceptions vary greatly among them. This lack of unity is obvious with regard to the kind of normative judgments over politics it has supported among the different "contractarian" philosophers and political thinkers. Positions have ranged from justifying royal absolutism (Thomas Hobbes, 17th century) to leaning on a republican understanding of just politics (Rousseau, 18th century) or accommodating a "liberal" philosophy (Rawls, 20th century).

Although one cannot disregard the fundamental differences and oppositions that appear between contractarian philosophies, it would be a serious mistake to ignore their underlying unity. This is not to be found in what they advocate politically but rather in a relatively homogeneous pattern of ideas. The most constant one is an inquiry into the legitimacy of powers. These are no longer considered as imposed by nature or by God. Contractarian theories provide another explanation: We live in a collectivity because we want it. We obey laws because we want them. Society and authority are the reflection of our will, if only implicitly. Legitimate powers are those that respect the conditions of this will.

This does not mean that obeying laws and respecting more generally the constraints that any social life carries are dependent on the fantasy of its members or that we can be free from our obligations "at will." Here the theme of the contract is precisely to signify that there is no contradiction between binding obligations and the personal will to enter them. We are bound to obey the rules in society, just as in any contract we engage willingly to perform services and duties that we are not free to disengage from without a penalty.

The emergence of social contract theories at the end of the 16th century should be understood in its historical context as the progressive awareness of a deep contradiction regarding the modern condition of politics. How is it that human beings both accept and reject authority, as shown by the bloodshed of the religious wars on the continent and the civil war in England? How is it possible to have people live together, while wars and revolutions give the constant example of what the philosopher Immanuel Kant (1724–1804) calls, in his 1784 *Idea for a Universal History From a Cosmopolitan Point of View* (fourth thesis), "the unsociable sociability of man"? Only a sound understanding of what human beings should accept in political matters—in other words, only a rational account of the legitimacy of power—could provide understanding and perhaps a solution to the evils of the time and grant a more stable future. This is ultimately the promise of contract theories.

The historical situation is not in itself the sole cause of such interrogations. After all, the course of history has always been a run of blood and misery. But the wakening of modern times makes it more intolerable *intellectually*. More and more, a new spirit inherited from the Renaissance values reason as a tool for improvement and insists on the equality in reason of human beings. This new sense of progress clashes with the dark situation of Europe in the 17th century and the harshness of social existence in the 18th. The new sense of equality is at odds with the hierarchical orders of the times. To these contradictions, contract theories have been the very much needed *intellectual response*.

The sense of novelty in contract theories is expressed in the constant rejection of conceptions of political life that we can label as "traditional" and that considered society as natural. No one more explicitly than Hobbes (1588–1679) in the first pages of his *De Cive* (*The Citizen*) has rejected this interpretation of society, inherited from Aristotle (and more generally from the whole of antiquity), for whom—man being a "sociable animal"—any human being seeks association with others in a city. For Hobbes, it is a delusion or a lie:

> The greatest part of those men who have written ought concerning Commonwealths, either suppose, or require us, or beg of us to believe, That Man is a Creature born fit for Society [. . .] which Axiom, though received by most, is yet certainly False. (I, chap. 1, para. 2)

Hobbes thinks that the prominence of strife and conflicts over benevolence and enjoyment in human ties belies the idea that society is about cooperation. He is drawn to the conclusion, which is shared by the whole contractarian tradition, is that if people live together, it is not out of a natural instinct. It is because they commit themselves to overcome the natural difficulty of togetherness. This commitment takes the form of a voluntary contract at the root of social existence.

History

It has sometimes been argued that the ancient Greeks and Romans recognized some form of social contract. It is difficult to concur with this view. The idea that cities had to start at some point with the gathering of a handful people was not foreign to ancient philosophers, it is true. For instance, Plato (ca. 427–ca. 347 BCE), in the third book of his *Republic*, relates the beginning of a city. But, he insists, it is the consequence of the fact that "human beings aren't self-sufficient" (369b); like other philosophers in antiquity, he pictures the beginning of societies (or rather cities) as the prime accomplishment of human nature, which would be incomplete otherwise. Contractarian theorists have a radically opposite point of view. For them, human beings are initially self-sufficient. However, this autonomy leads to unbearable conflicts and threatens their survival. Rather than being their natural destiny, society is their second-best choice (through a contract) to avoid extinction. This emphasis among contractarian theorists on the individual is another incompatibility of ancient theories for which the individualization of behaviors is the end of a well-formed city; as Plato sees it, for instance, the path to tyranny starts when "everyone has the right to do as he chooses" (*Republic*, 557b). In contractarian philosophy, the individual freedom of choice is the very basis of the contract that gives way to collective existence.

It is not necessary to multiply quotations and examples to realize that no Greek or Latin author would go so far as to consider that society was based on a contract between its members. For them, society was part of the natural order of the universe and nature provided a spontaneous impetus to collective life. Even when the idea is developed that social life is organized by conventions among human beings—an important claim by the sophists that is vigorously contested by authors such as Plato and Aristotle (384–322 BCE)—the role of these conventions is not quite similar to a proper social contract. The sophists think that conventions *organize* the collective life of human beings. Unlike contractarian theorists centuries after them, they never consider that they could have established it. Moreover, for the sophists, social conventions are mostly arbitrary or delusional, while for contractarian theorists, the convention between the members of a society is a rational one, based on the ability of human beings to make informed decisions and bind themselves through promises.

If we are to consider the influence of past theories on the emergence of contractarian theories, the legitimate reference should be not to antiquity but to social and political practices that developed later with the organization of feudal societies, which had, despite their own diversity, one remarkable characteristic: At its core, feudalism is the contract between a vassal and his lord. This contract states reciprocal obligations: the vassal serves, the lord protects. Eventually, feudal societies become shaped by a tight net of reciprocal obligations properly chartered as contracts.

This conception has led to spectacular practices, well beyond the Middle Ages. For instance, at the highest level of politics, contractualization was used in the 16th century by Henry VIII and Elizabeth I through a policy called "surrender and regrant": By contract, the rebel English lords in Ireland had to surrender their lands to the crown and receive them back in exchange for their submission, obtaining new extended rights of property they did not possess previously. This very idea will find an outstanding philosophical extension in Rousseau's *Social Contract:* According to him (as it is for John Locke before him and Kant after), the effect of the original contract is precisely to turn factual possession into legitimate property with the agreement of all.

Another (famous) example, the Mayflower Compact, showcases the continued taste in the 17th century for contractualization as a governance tool, even at a microlevel. Less than 2 decades after the death of Elizabeth, a group of settlers navigate to the new world—they later become the American icon of the colonization of the continent under the name "the Pilgrim Fathers." On the ship, the *Mayflower*, dissent provokes a serious crisis. To solve it, the passengers decide to "Covenant and Combine [them]selves together into a Civil Body Politic." This agreement has quite naturally been the subject of many interpretations, which see in it a sort of social contract, if not the "original American social contract." Such a comparison is flawed with misinterpretation. If a text such as the Mayflower Compact is to be regarded as having some tie to contract theory, it is not as an illustration and still less as an actualization of it. It is rather a good empirical example showing how the theme of the social contract is a highly sophisticated intellectual generalization of a rather common practice of the time to establish covenants in political matters.

It does not follow that the evolution from the practices of contract in the Middle Ages to contractarian doctrines is seamless, and even at its limited level the Mayflower Compact points out one of the most important differences. Medieval conceptions were rooted in hierarchies, whereas the Compact is an agreement between equal parties. Even if it is an exaggeration to describe the episode—as it has often been the case—as "democratic," it is nevertheless an example of the greater sense of equality in the 17th century. This sense is pervasive in contractarian philosophies and makes them break free from the very meaning of feudal contractualization, the goal of which was to regulate a constant exchange of services between the inferior and the superior. Contract theories, on the contrary, state that inequalities can be understood only *after* a primary agreement to live in society.

Contract theories are radically distinct from medieval conceptions on another crucial instance: While politics and their representation in the

Middle Ages gravitate around religion, contract philosophies are inherently secular. In the Middle Ages, religion is the ultimate justification of the authority of the princes whose power is regarded as proceeding from God's will. References to religion are repetitive, insistent, and lean on an often quoted passage of Paul's *Letter to the Romans* (XIII 1–2): "For there is no power but of God; and the powers that be are ordained of God. Therefore he that resisteth the power, withstandeth the ordinance of God: and they that withstand shall receive to themselves judgment."

Even if contractarian thinkers such as Samuel von Pufendorf (1632–1694) or John Locke (1632–1704) consider that the idea of a fundamental contract coheres with the plans of God, they never recognize a direct divine influence. The idea that society is the outcome of a covenant between human beings implies stepping away from such an immediate consecration of rulers by God. This is why, despite prudent maneuvers to avoid the accusation, contractarian philosophers have often been accused of being hostile to religion: Whether they adopt, like Hobbes, a strategy of compulsive justification of their positions by references taken from the Bible or whether they contemptuously dismiss faith as a source of comprehension for society, like Rousseau, they were equally considered in their time as impious.

Thus, the real novelty of contractarian theories is not that they assume the existence of a contract between the government and its subjects—this is the traditional aspect, inherited from the Middle Ages. The novelty is that they describe the very existence of society, the mere fact of living together, as the outcome of a pact between the members of society without direct reference to God's will. Assuming the idea of a global pact at the foundation of society, contract theories have transformed what was a local, discrete agreement of submission into a general mechanism accounting for all basic social and political relations.

The State of Nature

The assumption of contract theories that societies were created by an initial agreement purports the consequence that there is a state before society, a state in which human beings live *without being tied by the requirements of social life*. In the 17th century, an expression is coined to describe this moment when human beings are not formally engaged in the bonds of society: "the state of nature." The use of the word *nature* to characterize what is "not social" is remarkable: It is a direct translation of the new conviction in the 17th century that politics oppose society and nature. This antagonism itself is complex but rests on a conviction shared by all the authors: Society is about restraining an unlimited thirst for freedom and license that dominates the state of nature. Hence, there is a fundamental contrast: On the one hand, man (few authors reflect on the status of women, usually considering that the division of genders is a natural one) is naturally free and wants to be infinitely free; on the other hand, this freedom is not sustainable and ultimately leads to anarchy. Society is about the rules and constraints that bind this initial liberty. For this reason, contractarian authors have a repressive conception of society generally—even if, along with Rousseau or Kant, they explain that this repression is true liberty, liberty within autonomous rules.

In this perspective, the state of nature serves as a way to explain this liberty that men have given up by entering into the ties of society. In its most common versions, it is described as an age only dominated by "natural law," in the new understanding that the expression has had since the 16th century. *Natural law* is the law that determines the behavior of human beings without reference to any positive norm. In his 1690 *Second Treatise of Civil Government*, Locke has summed up what natural law entails for most contractarian authors: "perfect freedom [for men] to order their actions, and dispose of their possessions and persons, as they think fit . . . without asking leave, or depending upon the will of any other man" (chap. 2, sec. 1).

Nothing but strength or opportunity seems to limit men in their endeavors. It leads to a situation of at least potential threat, and more surely to violence and to anarchy. As Hobbes puts it, the "state of nature" is a "state of war"—"the war of everyone against everyone," to recall a famous quote from his *Leviathan*, published in 1651 (I 4). The place of strife in the state of nature is not always as pregnant as in Hobbes's philosophy: Rousseau strongly opposes the idea, considering that conflict follows the invention of society—an ill-regulated society without a social contract—instead of preceding it. Locke thinks that while not being constant, conflict may pose a problem at any moment. In all these cases, despite deep oppositions about the nature of conflict and its presence in the state of nature, the theme of the contract surfaces as a response to a potential menace: At some point, human beings decide to live together under the rules of a centralized government that will protect them as long as they limit their liberty and obey the rules that this government is in charge of enforcing.

John Milton's 1650 *Tenure of Kings and Magistrates* includes a good summary of these main ideas:

> All men naturally were borne [sic] free . . . and were . . . born to command and not to obey: and that they liv'd so. Till . . . falling among themselves to doe wrong and violence, and foreseeing that such courses must needs tend to the destruction of them all, they agreed by common league to bind each other from mutual injury, and joyntly to defend themselves against any that gave disturbance or opposition to such agreement. Hence came Citties, Townes and Commonwealths. And because no faith in all was found sufficiently binding, they saw it needfull to ordaine som authoritie, that might restrain by force and punishment what was violated against peace and common right.
>
> This authoritie and power of self-defence and preservation being originally and naturally in every one of them, and unitedly in them all, for ease, for order, and least each man should be his own partial Judge, they communicated and deriv'd either to one, whom for the eminence of his wisdom and integritie they chose above the rest.

Milton sums up the starting point of contractarian theories. Beginning with an anthropology that claims that the fate of man is liberty and commandment and not obedience, Milton points out the danger: If nobody obeys common rules, humanity is doomed to suffer from violence because there are no possible restraints on whatever action is undertaken by whoever wants something. However, the response is the ability of men to "foresee" the trouble and to surrender their original freedom and equality to secure peace against potential troublemakers. So is born the age of law, an age of limited freedom under law and authority. The substance of what is called "the social contract" is made of this agreement disclosing rights and duties, legitimacy and obedience.

Authors pay more or less attention to the detail of the state of nature. We find ample depictions in Hobbes, Pufendorf, or Locke; and Rousseau takes the idea of the state of nature to an unequal level of theoretical sophistication in his *Second Discourse* (published in 1755). In contemporary philosophy, the reprise of Sigmund Freud's narrative about primitive society in Carole Pateman's notion of sexual contract also gives a developed narrative about the equivalent of the state of nature. But authors such as Kant or Rawls mention only briefly the "state of nature." Its sole goal is to give a sense of what human attitudes toward engaging in political life could be if existing social constraints had no influence on the choices they make. Whatever its development in the different theories, the state of nature and its equivalent form an extraordinary hypothesis about a humanity freed from any kind of social constraint and depict, often under the guise of fiction, the experience of a humanity for which every path, from radical anarchy to complete social integration, is open.

Using the state of nature for a heuristic hypothesis about human freedom is not without problems. The first is its logics. If humanity has stepped away from nature, how can a proper understanding of the latter be retrieved since—as Rousseau put it—the necessary experiences to retrieve the "original man" are impossible to create? Some authors just assume, like Hobbes, that the state of nature can still be perceived in contemporary human behaviors: As soon as the constraints of rules stop making their grip felt, all the violence of nature comes back, with abuses, desire for violence, and strife. The second question is the one of the reality of the state of nature and, beyond, of the initial agreement that gave way to the existence of society. For Rousseau it is just a "conjecture," a state that is "no more" and that has "probably never been" (as Rousseau puts it in his *Second Discourse*). But Locke, for instance, never renounces the idea that the initial pact corresponds to some extent to historical events. Pufendorf—being careful not to contradict the biblical narrative of the creation of humanity—considers that "the whole human race has never at one and the same time been in the natural state" (Pufendorf, II, chap. 1): It is a fragmentary experience.

From Rousseau on, the interest in the historical accuracy of the contractarian narrative fades. Rousseau certainly indulges in a long, detailed, and forcefully evocative narrative about the state of nature in his *Second Discourse*. But he also decides to "lay facts aside, as they do not affect the question": The narrative of the state of nature is bound to remain a fiction. This renunciation to ground the theory of contract in empirical evidences is connected to the assumption that, as Kant puts it, the notion of a social contract is a pure a priori idea of reason. Kant opines that such ideas allow us to think and to understand reality but are not derived from experience. That is why—apart from an explicit fear of revolution—Kant criticizes empirical inquiries about the origins of the society: "It is vain to inquire into the historical origin of the political mechanism; for it is no longer possible to discover historically the point of time at which civil society took its beginning" (Kant, *Eternal Peace*, p. 146). Rawls's contemporary version of contractarianism follows this tradition: His own radical position advocates the suppression of particular information in the "original position," which is the first step to determine principles of justice, an undertaking that he presents as an analog of the social contract. Here, as with Kant, the reality described by the social contract is not the reality of *social facts* but the reality of a *social reasoning* giving way to an appropriate perspective on social reality. After Rousseau, contract theories are no more about history, they are a rational standpoint to study society.

An Individualistic Understanding of Political Life

Considering societies as owing their existence to an initial agreement involves three main questions that have been dealt with thoroughly throughout the history of contract theories: (1) about the status of the contracting parties, (2) about its motivation of the covenant, and (3) about its limit—who, why, what?

Answering the second question first is probably the easiest, because there is a certain consensus among authors about "why," at a certain point, human beings find it fit to engage in the ties of society, abandoning their initial freedom and equality to respect common laws and the authority in charge of enforcing them. All agree that the state of nature is by nature unstable and unsustainable: Unlike animals, the human fate is not to stay in nature. Only Rousseau disagrees with this view. He considers that it is only by accident, through some unknown catastrophe, that men have come to be in contact with one another, while they could have stayed forever in the animal state. This conception allows him to avoid the less gentle explanation that other philosophers were keen on offering: Man cannot remain an animal, because he is worse than the animals. As Pufendorf writes in *On the Duties of Man and Citizen*, man is "more wretched than that of any wild beast" (I 4) and "no animal is fiercer or more untameable than man, and more prone to vices capable of disturbing the peace of society" (V 6). Hobbes coins a famous sentence to sum up this somber anthropology in a vision of "the life of man, solitary, poore, nasty, brutish, and short" (*Leviathan*, chap. 13). Whether by accident (for Rousseau) or out of the continuous strife that the vices of men initiate when there are no rules to tame them, there is a moment when the natural state of man is untenable. The reason why man agrees on a contract that forces him to renounce his initial liberty for the sake of securing what is the most important to him (property for Locke, rights for Rousseau, life for Hobbes, a bit of everything for Pufendorf) is the direct consequence of this state of violence to which the initial condition has evolved.

The reason why a contract seems to be compulsory to contractarians shows how narrowly it is linked with a conception of humanity—in other words, an anthropology. The anthropology at stake has one consensual element among philosophers who, otherwise, have very different views on human nature: Man who is party to the contract is defined as an individual. Baruch Spinoza (1632–1677) spells it out in the *Tractatus Theologico-Politicus:* "Each individual transfers the whole of the power he has of himself to the community" (chap. 14). Hobbes uses the first person to individualize the engagement in the compact: "I authorise and give up my right of governing myself to this man, or to this assembly of men, on this condition;

that thou give up, thy right to him . . ." (*Leviathan*, II 17). Rousseau equally emphasizes the individualistic dimension of the social compact, which is for him "the total alienation of each associate, together with all his rights, to the whole community; for, in the first place, as each gives himself absolutely, the conditions are the same for all" (*On Social Contract*, I 6).

Such formulations clearly indicate the reorganization of the conception of society with contract doctrines: Collectivities are to be understood starting with the individuals. They also lead contractarian philosophers to develop anthropological inquiries in order to determine that what is genuinely human as the core opposition of contractarian theories is the one between a *state before society* and *civil society*. This anthropology, diverse thought the ideas of the authors can be, is based on some convictions that they all share. All agree, for instance, that man is naturally free, equal to any other. The question of reason is more debated, and usually authors describe humanity in the state of nature as dominated by desire and appetite, the development of reason requiring in their opinion exchanges made possible only by the existence of society. Once again, Spinoza provides an accurate summary of this anthropology in his *Tractatus Theologico-Politicus*:

> The natural right of every man therefore is determined by appetite and power, not by sound reason. For all are not constituted by nature to act according to the rules of reason. On the contrary, all are born ignorant of everything; and before they can know the true rule of life, and acquire virtuous habits, a great part of their lives must already have passed. (chap. 16)

However, the consequences drawn from such a description, which is not solely Spinoza's, but could be agreed on by most authors of the 17th- and 18th-century contractarian thinkers, are often at odds. Hobbes considers that the fear of death, and hence the potential threat that any human being represents to any other, is the dominant passion, while Rousseau notes that the fear of death is too complex an idea to have it attributed to the natural state of man. Locke insists on the role of property and considers that the need to protect possession is the driving factor for the establishment of a civil society. For many contractarian philosophers, while a person is described as an individual, this individual is often engaged in different kinds of social intercourses. These do not lead to a properly established society—these relations are not formalized as laws and enforced by the existence of a state. As Kant repetitively states it in *The Science of Right*, the natural state of man can be a "social state" even if it is by no means a "civil state" because the civil state is characterized by the establishment of permanent laws, the violation of which is punished by a recognized authority.

Only Rousseau seeks to build a philosophical anthropology that would be a radically individualistic one, by eliminating from his consideration of man before the social contract any type of interactions between human beings in the state of nature. It seems contradictory to him to establish society on the basis of an individual agreement if we are unable to consider the individuals as such, without connections to one another. Even Hobbes's "war of all against all" in the state of nature seems to him a contradiction: If one is on her or his own, why should there be a conflict and how could two individuals who are entirely free to flee fight if they have no possession to defend and no complex idea such as honor and recognition to drive them? This opposition is far from being a point of detail. It is a way for Rousseau to develop to its furthest consequence the individualistic standpoint of contractarian doctrines and allows him either to develop or reformulate them or to overcome some of their difficulties.

For instance, Rousseau's radically individualistic stance gives him the possibility to address an aspect of the social contract that is often carefully avoided by most authors: the question of women. Contractarian philosophers tend to think that female subordination is natural—for example, Pufendorf, for whom marriage is part of the natural law. Hence, women have hardly a word to say in the original covenant, and the social contract is indeed gendered—Pateman is far from being wrong in reading the social contract as a sexual one, opposing the social contract, which is about freedom, to the sexual contract, which is about subordination. Even if Rousseau's claims about the legitimate subjection of woman to man in society are conventional and quite disappointing for the modern reader, his initial intuition is remarkable.

An individualistic society is gender neutral. This claim was long forgotten before being rediscovered in the 20th century—for instance, by Rawls, according to whom the decision about the principles of justice has been made under a veil of ignorance so thick that one does not even know the gender of the founder.

The emphasis on the individual is also decisive for the substance of the contract—the "what question." With his radical insight, Rousseau is able to overcome some of the least liberal aspects of contractarian doctrines and to give them a more democratic undertone—though technically Rousseau was never in favor of democracy, a regime he imagined only fit for "a people of gods." Admitting a social relation of some sort among men before the founding contract of society means that there is something wrong with human beings in the state of nature since they are unable to cope with it in a sustainable fashion. For many contract philosophies, the place of the individual is purely disruptive: Individuals tend to free themselves from the natural law—the rules that commend the action of man in the state of nature—since there is no institution to guarantee retaliation for misdeeds. Punishment requires the invention of a developed society and is a function of the state that is created after the contract is agreed on. If individuals spontaneously behaved according to the law of nature, indeed, there would be no need for a social contract. To be precise, one of the goals of the contract is to have man abandon natural law to ensure what this natural law cannot guarantee: its own enforcement. The risk is then to interpret the contract that founds a society either in a purely repressive way—this is basically the case with Hobbes—or to be exposed to a fallacy, as might be the case with Locke.

Indeed, if the main issue in human interactions in the state of nature is the menace of violence, because individuals have the right to anything they wish under the natural law, the scope of the contract can be strictly limited to guarantee the absence of conflict by all means. It is the case with Hobbes, who considers that the only limitation of the authority to which the contract has transferred each individual's power and right is that it cannot harm the life of its subjects. In such a context, even if it has been argued that Hobbes was the first "liberal" thinker, authoritarian regimes can be perfectly legitimate ones.

But if, as Rousseau did, one starts with an individual isolated from the others, then conflict only appears when society is formed because it puts into contact with each other human beings who are not meant to be placed in such a situation. Hence, suppressing conflicts means retrieving the initial state by finding a way—which is precisely the contract—to guarantee what has been lost with the gathering of individuals who were initially separate entities. In a way, Rousseau plays the *civil state* against the *social state:* society without laws oppresses the individuals; society with laws should enable them to thrive again in a superior way. This leads to a major step ahead in the values linked with the idea of a social contract. For Rousseau, the aim of the initial contract is not only to protect the physical life of the individual but also to protect his or her inherent qualities: freedom, rights, and equal standing with everybody in society. To that extent, he follows the lesson of Locke in his *Second Treatise*, who says that "law, in its true notion, is not so much the limitation as the direction of a free and intelligent agent to his proper interest" (sec. 57). And Rousseau extends the idea to its ultimate consequence: if law is here to protect liberty, as Locke claims it, then Rousseau feels that the logical follow-up is that "obedience to a law which we prescribe to ourselves is liberty" (*On the Social Contract*, I 8).

Rousseau's difference with Locke is precisely that Locke also considers that human beings in the state of nature are not isolated. He goes so far as to think that they have some common understanding of the law (of nature) they should respect. But the problem with this view—and the possible fallacy in the reasoning—is that it shows that social relations are right away flawed with conflict and violations of norms that are however respected. The formation of civil society through the contract only adds government and law to the problems of society but does not solve them. By retracing the origins of society, not to a state of nature that would have been to some extent social but to the idea of the individual considered in himself or herself, Rousseau is in a better position to reinforce the claim of an original contract that is in place to preserve individual rights, because they can be thought independent from any kind of interactions.

Beyond the Contract

No author more clearly than Pufendorf in his *Duty of Man and Citizen* has grasped the essence of the questions raised by contractarian theories. "For a state to coalesce regularly," Pufendorf writes, it is necessary that first each individual agrees "to a permanent community" with the others, or, in other words, it is necessary "that they mutually desire to become fellow-citizens." This is the first compact. Then, once engaged in society, the collectivity has to decide in "one decree" under which regime it chooses to live. Then,

> another compact is needed, when the person, or persons, upon whom the government of the nascent state is conferred are established in authority. By this compact these bind themselves to take care of the common security and safety, the rest to yield them their obedience; and by it also all subject their own wills to the will of that person or persons. (II 6)

One compact is thus about living together, another about submission to authority. This theory, sometimes known as the "two-pacts theory," underlines the twofold character of contractarian theories: on the one hand, the question of power; on the other, the question of togetherness—in other terms, *the question of political legitimacy* and the *question of social diversity*. If Rousseau's formula "the social contract" has been so successful in summing up the inspiration of contract theory, it is precisely because the original contract is not solely about government. It is about the very existence of any collectivity, any *society*. The purpose of contract theory is thus not only to identify the characteristics of legitimate power. It also deals with the mere possibility of living together and raises the question of social ties.

We can safely say that this new question is born with contractarianism and that contract theory has been the first to give an account of it. With the emergence of the figure of the individual, social bonds appear frailer because one realizes the power that each member has to seclude himself from ties that can be challenged. While tyranny has always been considered as a political evil, contract theories also admit that, beyond the oppression, seeing communities torn down by internal disputes is an equal danger. The times of revolutions and civil wars that saw the birth of contractarian doctrines gave ample illustration to this fact to authors who tried to figure out what response to the tragedies of the time could be offered.

Pufendorf, by separating the will to live together from the consent to power, spells out this new condition—the social condition—of politics in the wake of modernity. However, if he is analytically right, he has been seen as politically wrong, and the rest of the contractarian tradition has rejected the theory of the two pacts. The fundamental problem of modern politics might be twofold—living together, obeying a legitimate authority—but it does not consist of two separate problems. Determining the right authority, the right "principle of justice," as Rawls puts it in the 20th century, is inseparable from our way of living together. According to Rawls, it is not the case that there is our political existence on one side and the kind of relations we have with other members of our society on the other side. Politics, relations of authority, pervade the entire network of social links: Political sociology always meshes with the framework of governance. With contractarian doctrines, political philosophy has entered the age of ideology: Social relations are to be accounted for in terms of politics.

By the same token, the presence of contract theory in political thought has been abating whenever those questions have appeared as secondary or even suspect ones or when alternative theories of modernity have shown a better way of giving explanations of the facts that contract theorists wanted to explain. Another kind of critique has also come from those who, as David Hume (1711–1776) puts it in one of his essays dedicated to the critique of contractarian theories, "trace up" "government to the Deity." The very inspiration of a secular doctrine is for these critics its very and most fundamental flaw. But as this kind of normative position has lost its grip on Western political thought, the real critique has come from doctrines that have assumed their own identification with modernity, showing that contract theory could embrace the whole of modern politics.

In the first phase, after the French Revolution, one of the major problems of contract theories—social plurality and the question of how to live together in society—loses some of its importance. The 19th century was an age of nation building and social reengineering. The national imaginary favors the conception of tightly knitted communities, whose main problem was to get rid of external influence. The neighbor was not a problem provided that he or she was from the same kin, and the legitimate government was not primarily the one that grants rights but the national one. Socialist doctrines, then so influential and whose influence extended far into the 20th century, were equally at odds with this question of social plurality. Their insistence on collective organization could not but repel the general inspiration of contract philosophies. Insisting on class conscience, on class organization, meant having very little interest in cultural diversity or in political struggles that would not be linked to structural, economic factors. It also meant a clear inability to even consider the significance of any individualistic point of view on society, which contract theories always involve. Eventually, the style of political thought involved in contractarian theories also could be challenged with the development of a scientific approach of politics favoring the collection and the analysis of empirical data: It did not leave much room for theories that were highly speculative.

Another type of critique of contractarian doctrines deals with the issue of liberty. One of the most constant representations of liberty in contractarian theories rests in its dramatic character: Human beings, left on their own without the restraints of a political order, are led into a state of violence. But since the 18th century, liberal and then libertarian philosophies have considered that the lack of liberty is the sole source of violence. Left to themselves, human beings are able to live in harmony. This is the gist of the inspiration of market theories, which have a reach that is initially far beyond a limited explanation of how economics work or should work. To be free, only local contracts are necessary, and the government, far from being the outcome of a collective decision to protect liberty, should not intervene in matters regarding liberty. Safeguarding liberty does not need the establishment of an original contract but freedom from the government.

These competing doctrines also have shown their limits, again placing contractarian doctrines in the foreground. The publication in 1971 of John Rawls's *A Theory of Justice* is largely recognized as a milestone in the renewal of interest in contractarian approaches. It coincided with the awareness that a free market could not be a substitute for a just society, and that questions of principles were raised through new political issues such as discrimination among a population: national citizenship alone was not enough to build a political community. The rise in the United States and the United Kingdom of political agendas more centered on individuals in social interactions than in the previous decades, as well as the declining influence of socialist and Marxist interpretations of society, also allowed the contractarian framework to appear as a credible alternative. They have provided, under a new guise, a way to conciliate the renewed prominence of the figure of the individual and the continued necessity of organizing collective existence. In societies that are becoming more and more diverse, and where governance faces repetitive crisis, contractarian theories might not offer the solution but should still count among the most powerful tools we have to understand the contemporary world of politics. This is indeed the problem of what we have eventually called "democracy" and where we now recognize the very definition of our age.

Thierry Leterre
Miami University John E. Dolibois Center
Luxembourg

See also Conflicts; Hobbes, Thomas; Individualism; Locke, John; Natural Law; Rational Choice; Rousseau, Jean-Jacques

Further Readings

Barker, E. (1980). *Social contract: Essays by Locke, Hume, and Rousseau.* Westport, CT: Greenwood Press.

Barry, B. M. (1989). *Theories of justice.* Berkeley: University of California Press.

Bawden, D. L. (1984). *The social contract revisited: Aims and outcomes of President Reagan's social welfare policy.* Washington, DC: Urban Institute Press.

Boucher, D., & Kelly, P. (1994). *The social contract from Hobbes to Rawls.* London: Routledge.

Hobbes, T. (1983). *Philosophical rudiments concerning government and society* (1st ed.; H. Warrender, Ed.). Oxford, UK: Clarendon Press. (Original work published 1641, as *De Cive*)

Hobbes, T. (2002). *Leviathan* (A. Martinich, Ed.). Peterborough, ON, Canada: Broadview Press. (Original work published 1651)

Kant, I. (1914). *Eternal peace and other international essays* (E. D. Meade, Ed.; W. Hastie, Trans.). Boston: World Peace Foundation. (Original work published 1795)
Kant, I. (1997). *Science of right* (W. Hastie, Trans.). Edinburgh, UK: T&T Clark. (Original work published 1790)
Lessnoff, M. (1986). *Social contract.* London: Macmillan.
Locke, J. (1982). *Second treatise of government* (R. H. Cox, Ed.). Arlington Heights, IL: H. Davidson.
Milton, J. (1649). *The tenure of kings and magistrates.* London: Matthew Simons.
Morris, C. W. (1999). *The social contract theorists: Critical essays on Hobbes, Locke, and Rousseau.* Lanham, MD: Rowman & Littlefield.
Pateman, C. (1988). *The sexual contract.* Palo Alto, CA: Stanford University Press.
Pufendorf, S. (1991). *Pufendorf: The two books on the duty of man and citizen according to natural law* (Cambridge texts in the history of political thought). Cambridge, UK: Cambridge University Press.
Rawls, J. (1999). *A theory of justice.* Cambridge, MA: Belknap Press of Harvard University Press.
Rousseau, J.-J. (1997). *The social contract and other later political writings* (V. Gourevitch, Ed. & Trans.). Cambridge, UK: Cambridge University Press.
Saastamoinen, K. (2010). Pufendorf on natural equality, human dignity, and self-esteem. *Journal of the History of Ideas, 71*(1), 39–62.
Spinoza, B. (1670). *Tractatus theologico-politicus.* Retrieved November 2, 2010, from http://www.yesselman.com/ttpelws1.htm

COOPERATION

Many definitions of politics emphasize the conflictual aspect of human relations. But cooperation should be considered the essential element of politics even if it is not always achieved at the degree that would produce optimum results. Cooperation is action for the common benefit. Only human beings are able to cooperate and abide by collective rules for their common interests, including making exchanges in their mutual benefit, forming coalitions and stable organizations, making enforceable decisions on collective affairs, and living in large communities under shared norms. This entry describes some of the aspects of human existence in which cooperation can play a key role, examines ways in which game theory can illuminate the nature of cooperation, and discusses some of the factors that are important in developing and maintaining cooperation.

Crucial Areas for Cooperation

All fundamental problems in politics face the crucial question of how, under what circumstances, and to what extent human beings can be motivated to cooperate in their common interest. Cooperation is at the core of the issues of conviviality, democracy, peaceful coexistence between different communities, and the preservation of human life on Earth, as is briefly reviewed in the following paragraphs.

Community

In what certain classical authors called "the state of nature," conflict is pervasive. If human interactions are unconstrained, anyone, with the advantage of surprise, can try to impose his or her will on others. But if all do, then people may find themselves living in a state of chaos in which, in Thomas Hobbes's famous words, life tends to be "solitary, poor, nasty, brutish, and short." In such an environment, it is not reasonable to risk unilateral cooperation, while cooperation within groups is precarious. However, human beings can do better. People can agree on creating a government equipped with tools of coercion to enforce rules mandating cooperative actions that produce beneficial results for society as a whole. The government may apply sanctions against "defectors"—that is, violators of mutually beneficial rules of conduct—discourage free riding on public goods, and craft incentives for cooperation. People can rationally accept conditional consent. By an agreed "social contract," the efficient outcome of civilization or "commonwealth," in which each can live in peace and security, can be attained.

Democratization

In situations of institutional regime crisis in which authoritarian rulers cannot go on as they were accustomed to do, actors with opposite political regime preferences can generate violent conflict or a civil war in which both sides may fight to eliminate each other. Eventually, one of the sides can become a single, absolute winner. But choosing confrontation, with an uncertain outcome, also entails the risk of becoming an absolute loser, as well as the costs of significant destruction on both sides. In contrast, by anticipating the foreseeable consequences of their choices, either the rulers or the opposition leaders can offer conditional, retractable cooperation. Negotiations can lead to a provisional compromise, including the calling of a multiparty election that does not secure an absolute winner, which may open further developments in favor of either of the actors involved, as has happened in so many cases of democratization in different parts of the world since the last quarter of the 20th century.

Deterrence

International relations have traditionally been dominated by conflict, which culminated in the Cold War between the United States and the Soviet Union during the second half of the 20th century. However, a major clash was prevented by underlying cooperation. An arms race triggered by the Soviets choosing to build new weapons and the United States doing the same, and vice versa, put both countries at risk. The "balance of terror" without an actual frontal war was durable because, with nuclear weapons and the possibility of total destruction, the stakes were so high. The United States and the Soviet Union attempted to cooperate in order to reduce and control arms through negotiations and agreements. Nevertheless, the arms race stopped only with the dissolution of one of the players.

Environment

Climate change has become a broad concern, leading many people to call to stop the planet from overheating. Yet some skepticism persists regarding human agency in climate change, and few governments are willing to deal with the problem by themselves. The Kyoto Protocol, placing voluntary limits on emissions of greenhouse gases, was formally accepted by 36 developed countries for the period 2005 to 2012. However, the United States and Australia initially refused to sign. Even some of the protocol signatories, as well as many developing countries, including most prominently China, continue to tolerate high levels of industrial emissions known to exacerbate the greenhouse effect. Without international cooperation, the world may be condemned to an eventual burning, according to those who hold the global warming hypothesis. Most countries seem to have recently found some incentive to avoid being sanctioned for their misconduct. The Asia-Pacific Partnership on Clean Development and Climate has attained new cooperation on development and technology transfer enabling a reduction in gas emissions. The United Nations held successful talks with virtually all countries in the world to replace the Kyoto Protocol after its deadline expires. If things get bad enough, then cooperation may flourish.

Developing and Sustaining Cooperation

The conditions in which cooperation among different actors for their common interest can emerge and hold up, such as in the cases mentioned above, can be illuminated with the help of some analytical tools provided by game theory. In particular, competitive, non–zero-sum games, such as the famous "prisoners' dilemma," involve different combinations of cooperation and conflict. When mutual cooperation can produce gains for all participants, it is said that the sum of the players' benefits is positive. But people may fail to cooperate with others even if cooperation would produce a better collective outcome for all the participants.

The lessons from prisoners' dilemma can be applied to any group or community facing a cooperation problem among its members. On the one hand, the previously conditions of the state of nature, civil war, arms race, or destruction of natural resources can be conceived as inefficient, equilibrium outcomes of prisoners' dilemma type of interactions. On the other hand, the state of civilization, democracy, peaceful coexistence, and conservation of the atmosphere can correspond to alternative, efficient, although somewhat vulnerable, outcomes of this type of game produced by mutual cooperation.

Game theory models with "repeated" games suggest that the greater the uncertainty over the future relationship and the higher the number of interactions between the parties, the more conflict can diminish and mutual cooperation emerge and be sustained. If people are going to engage in repeated interactions, it may make sense to try to cooperate in order to receive others' cooperation in the future. A community or institutional setting in which everybody can expect to keep interacting with the same people regularly for some time in the future may include a household or a neighborhood, the workplace, a mall or a school, a professional organization, or a political party, as well as the city, the state, the empire, or the world one is aware of living in and where one intends to stay. A reasonable behavior for repeated interactions with other individuals with some common interest can imply conditional cooperation and a positive response to the other's behavior. In the long term, cooperation may spread and become the prevailing way of conduct.

According to this insight, we can observe that cooperation, as can be substantiated in the form of collective action or a joint organization, is indeed more intense and sustained among certain groups of people interacting for long periods of indefinite length. Cooperation should be higher, for example, among members of a condominium association rather than among motel clients, among town residents rather than among tourists or occasional visitors, among fixed employees rather than among temporarily unemployed people expecting to find a job soon, among civil servants rather than among seasonal workers, and among store owners in a commercial mall rather than among sporadic vendors in a street market.

If people reciprocate to cooperation with cooperation in repeated interactions, they can build a good reputation for themselves that may move others to cooperate with them. Feelings of trust may emerge among people having information about others' past actions and among new participants obtaining regular positive retribution for their conduct. In the middle or long term, increasing and sustained cooperation among members of a community may induce them to construct institutional environments that limit individual competition and tend to homogenize the population. Internal sanctions against defectors can go together with the promotion of values such as honesty and empathy with the distress of others, thus reinforcing social cooperativeness. Emerging and self-sustaining cooperation in the long term can also help promote conviviality, democracy, peaceful coexistence, and environmental conservation mentioned earlier.

Nevertheless, the development of cooperation among members of a community requires some mutual commitment to stay within its contours. If, conversely, people living within the same institutional setting consider themselves as belonging to different groups with opposite goals, asymmetric relations of cooperation and conflict within the community can develop. The difference between groups can be based on language, race, ethnicity, religion, family or tribal traditions, contrary economic interests, or adversarial preferences for the location of public goods. Thus, high degrees of success or failure in attaining cooperation within a group do not necessarily correspond to socially efficient or inefficient solutions. The collective strength of some groups may indeed provoke conflict with other groups or favor asymmetric and biased redistributions of resources, hindering more satisfactory outcomes for greater numbers of people.

Josep M. Colomer
Universitat Pompeu Fabra
Barcelona, Spain

See also Coalitions; Competition, Political; Conflict Resolution; Contract Theory; Coordination; Democratization; Deterrence; Game Theory; Prisoners' Dilemma

Further Readings

Axelrod, R. (2006). *The evolution of cooperation*. New York: Perseus. (Original work published 1984)

Poundstone, W. (1993). *Prisoner's dilemma*. New York: Doubleday.

Riddley, M. (1996). *The origins of virtue: Human instincts and the evolution of cooperation*. London: Penguin.

Schelling, T. (1981). *The strategy of conflict*. Cambridge, MA: Harvard University Press. (Original work published 1960)

Taylor, M. (1987). *The possibility of cooperation*. Cambridge, UK: Cambridge University Press.

Coordination

An initial statement of what is typically meant by the pursuit of coordination is the ambition to achieve greater coherence among the departments, agencies, authorities, services, and professions of government, as well as those providing public services under contract or license, each of which has some responsibilities for a given problem or issue. This entry provides a framework for understanding coordination in terms of the problems it seeks to address and the means through which its goals are pursued. It also discusses the way in which the concern with coordination evolved, the main theoretical approaches that have been taken in research on it, and the limitations of efforts to provide evidence of its effectiveness.

Goals and Means

The word *coordination* is sometimes used to describe a class of policy goals or outcomes, loosely labeled as "coherence." This entry distinguishes several distinct goals that fall under this general term. Each of these goals, though, consists to some degree in the reduction or even removal of tensions among policies, so that they do not undermine each other and, perhaps, instead even support each other.

Because coordination covers a range of initiatives designed to combat certain types of problems, one way to understand its meaning is as something offered as the opposite of any or several of those problems. Broadly, the problems for which greater coordination is presented as a solution can be divided into six kinds, ranked here in order of declining ferocity of conflict between the imperatives set out in the institutional or organizational guidelines, organizational or professional goals, incentives, regulatory requirements, and legal duties (see Christopher Pollitt, 2003):

Conflict. Mutually undermining imperatives lead to conflict. For example, goals of promoting economic growth and limiting the environmental impact of human activity quickly come into conflict, but almost every government is expected to promote both to some degree, and they are often given to distinct agencies to pursue. Another such case arises when one regulatory agency demands adaptations to buildings to enable wheelchair access, while another regulatory body forbids it on the ground that they would violate the integrity of a historic building.

Contradiction. This entails inconsistent imperatives that have not actually come into direct conflict but may do so. For example, consider the goal of designing welfare systems to get people who have few skills and some disadvantages, such as learning difficulties or mental health problems, into the work force. Such a program may conflict with the imperative of productivity even in a strong economy, but the conflict may remain latent. However, in a weak economy, the inconsistency in goals may become important and constitute an actual conflict.

Competition. This involves imperatives that are not strictly inconsistent but, in practice, come into conflict or at least are in rivalry for priority or resources. This is a perennial problem, in which the competition between emphasizing investment in secondary and tertiary education and between crime reduction and reduction of the fear of crime can serve as examples.

Incoherence and Fragmentation. Imperatives that are neither mutually contradictory nor in competition may have little to do with each other, when arguably gains could be achieved if they were brought together. The frequent complaints of "departmentalism," "silo mentality," and "territorialism" between neighboring local authorities are all examples of this kind of problem.

Duplication. This refers to similar or even identical imperatives expressed in different arrangements that result in wasteful and excessive use of resources. For example, there are often cases where very similar programs are provided by both health authorities and local authorities or by local authorities and probation services; a similar issue can arise when multiple regulators all conduct on-site inspections, each looking at the same sets of organizational records or processes but for different purposes, where there might be possibilities to reduce the burden on regulated bodies by combining inspections.

Pursuit of coordination as a goal rests on the claim that better outcomes such as health, educational achievement, security against crime, economic development, or effectiveness in defense are better promoted when the policies pursued by each particular agency, department, or authority within the system of governance stand in little tension with each other. To many people, this claim seems obviously and intuitively true. However, as we shall see below, it is contested, and some of the arguments offered against it have very distinguished pedigrees in both organizational and political thought.

Often, however, "coordination" is also used to describe the *means* by which the aims of better integrated policy are pursued. Indeed, in this respect, the term is used to refer to

- general *processes*, such as markets, hierarchy, clans, or clusters, as broad types of change process that bring people and organizations into some kind of alignment with and in adjustment to each other;
- policy *instruments*, such as incentives, directives, or the provision of information (or, more simply, sticks, carrots, and sermons, which form the canonical trinity of types of tools of government); or
- very specific interorganizational *arrangements*, such as joint committees, joint budgets, or even activities or outputs such as a "one-stop shop" or shared services centers that are outlets to the service users or the public.

For the most part, governments use these means in combination. Rarely do they rely exclusively on one type of process or instrument even in a single field of policy, let alone across the whole span of public management. Much of the recent effort in the study of coordination is now devoted to the effort to understand the likely consequences, in different circumstances, of different combinations of processes, instruments, and arrangements. A common slogan for this is that "it's the mix that matters."

Just what relationship these means of coordination have to the goals of coordination is the subject of some disagreement and some disappointment. For example, some researchers have found that the very proliferation of large numbers of joint committees, budgets, and customer service centers itself produces new kinds of fragmentation. Certainly, all these things are expensive to establish and run. This raises the question of whether the gains in improved policy outcomes are worth the higher administrative costs of introducing and operating all the coordinating machinery. There is also evidence from the study of mergers that directing managerial attention to the working of committees and budgets and centers can distract them, sometimes for quite long periods, from the larger policy goals, while they concentrate on the minutiae of interorganizational relations. It would be quite wrong to say that coordination efforts typically fail or cost more than they are worth. Indeed, the literature includes many positive evaluations. But as yet, we lack sufficient evidence to make many general claims about when, where, and why the means support the goals and to what degree they can be expected to do so.

Origins

Coordination has been an eternal concern of government and the public services. Between the different branches of the armed forces, it has been a topic for institutional innovation, political anxiety, and intellectual work since antiquity. The need to reduce conflict between the impacts of the tax-collecting and economic activity–stimulating work of different functions of government was a matter of enduring concern throughout the Roman Empire. Incoherence and conflicting jurisdiction became a major problem for the absolutist monarchs of early-modern Europe, because the combination of the sale of offices, the farming of taxes, and the conflict of monarchical laws with local systems of laws produced growing contradictions.

With the development of responsibilities for the welfare of citizens in domestic policy toward the late 19th and early 20th centuries, the issue came to be of central importance in many countries. The "Progressive Era" reforms in late-19th-century United States produced some systematic thought on the question. In British imperial public administration and especially in India after 1857, the need for coordination became a constant litany.

In 1918 in Britain, the Haldane Committee published its report, examining more systematically than any previous study had done the various principles and administrative designs available by which greater coordination might be pursued. Options considered included those of organizing around geographical jurisdiction, clientele, or function, each requiring different approaches to horizontal interorganizational linkage.

In the decades after World War II, many countries experimented with mergers to produce ever larger departments or groups of departments in the hope that this would produce greater cohesion, but the results were generally disappointing. In the

1960s, some came to pin their hopes for coordination on techniques of central control rather than in structural change, emphasizing management by objectives, program analysis, and review and other forms of managerial and financial accountability. Yet by the mid-1970s, disappointment with these too had become widespread.

The 1980s and 1990s were characterized in Britain, New Zealand, Australia, and the United States by a series of reforms, closely documented in the scholarly literature, which are commonly labeled *the new public management*, although this term probably gives them greater unity than they in fact exhibited. By the mid- to late 1990s, concern was growing that these reforms had overemphasized the role of the dedicated specialist agency, often working under a contract for the performance of its own principal function exclusively, to the detriment of coordination. In Britain, the aspiration for a major initiative at every level toward greater cross-government working was given the label "joined-up government" by the New Labor administration that came to power in 1997; in practice, its interest in this waned significantly after a few years. In Australia, the preferred term was the *whole of government* approach; political commitment there was sustained for a longer period. The movement had no national programmatic emphasis in the United States, not least because many of the functions that would be candidates for spearheading greater coordination are in that country administered at state level. Indeed, it was at the state level that a wide range of efforts in things described variously as integration, horizontal working, and "network governance" came to be pursued.

In practice, different reform programs have pursued rather different things in the name of coherence and coordination. In some countries, the priorities have been around integrating access to specific local services targeting the same clienteles, through "one-stop shop" centers on main business streets or online. In others, the priority has been the formation of joint boards between a modest number of particular public authorities to manage special programs. In yet others, the principal effort has been put into designing systems of government-wide accountability to sharpen incentives for the pursuit of particular outcomes, irrespective of agency; finally, for some governments, improving the mutual understanding and joint efforts of their civil servants working in different departments of state to formulate policies is the greater priority.

It seems that the high watermark of this most recent tide of efforts in coordination had passed by the late 2000s, as governments became more interested in "modernization," supporting individual service user choice of provider, or as fiscal positions deteriorated in the wake of the financial and then the general economic crisis. No doubt, however, in due course, interest in aspirations for enhanced coordination will revive.

Although some advocates in the late 1990s had perhaps hoped otherwise, few votes have been won on the basis of greater coordination. At best, the contribution to electoral success achieved by coordination is indirect. Moreover, many of those who make the greatest use of those frontline public services for which coordination is typically a high priority are often poor and, therefore, typically less likely to vote. But this measure understates the political importance of the issue. Coordination has provided governments—and in particular, center-left governments—with an aspiration and ideas for programs of reform to the machinery of government with which to answer the charge that government is somehow inherently incapable of improving its performance. Although the details of administrative arrangements are of little interest to most voters, the ideological battle between those who hope for better government and those who argue for less government is a central one in defining political projects, and it is to this conflict that the debates about coordination make their political contribution.

Theories

Four kinds of questions about coordination have been the principal subjects for theoretical work in the literature:

1. explaining the waxing and waning over time and the differences between countries in political commitment to put into action programs of reform for coordination;
2. explaining the emergence, whether by conscious and deliberate choice or otherwise, of forms of interorganizational relations that have prevailed in pursuit of coordination in different countries, policy fields, and so on;

3. explaining the choice of instrumentation made by particular governments in pursuit of coordination; and
4. explaining the conditions under which we can expect certain programs to be successfully put in place administratively, or—more ambitiously—to make a positive difference to outcomes.

On the first and second questions, academic writing about coordination is a field where empirical work dominates both theory development and theory testing. Certainly, the grand explanatory frameworks of rational choice, historical institutionalism, structural functionalism, ideationalism, and so forth that dominate much of political science and international relations research and from which particular theories are often derived are not addressed by many scholars working on coordination.

The most commonly deployed argument in answer to each of the four questions is the broadly historical institutionalist one that the prior institutionalization of departments, organizations, tiers, and functions among the public services continues to inhibit the scope for greater coordination and joint working across boundaries. Much of the literature is devoted to the identification of institutionalized barriers to cooperation. Making lists of barriers, obstacles, constraints, or, conversely, facilitating conditions abounds. What is often unclear about the theoretical status and role of these sorts of factors is whether they are offered as necessary conditions for successful implementation of any particular level and degree of coordination, jointly necessary conditions for it to be sought at all, or simply factors that influence probabilities. Nor is it very often clear what causal relationship these structural factors have with agency. A partial exception is the attempt by Perri 6, Diana Leat, Kimberly Seltzer, and Stoker Gerry (2002) to apply the neo-Durkheimian institutional theoretic tradition to develop theories about both the first and the second questions.

On the first question, 6 and colleagues argue that political commitment to coordination is sustained principally during periods of centrist government and that this pattern can be traced back throughout the 20th century. By contrast, Helen Sullivan and Chris Skelcher argue that it is a recent development and an artifact of reaction against the excesses of the reforms of the 1980s and 1990s commonly referred to as the "new public management." The disagreement here turns on empirical questions about just how different the recent efforts in coordination really are from those pursued in previous periods of administrative history. Whereas 6 and colleagues and Christopher Hood see continuities in the elementary forms of organization and interorganizational relations, Sullivan and Skelcher argue that the range of tasks with which the public services are charged now and the context of the reaction against the excesses of the reforms of the 1980s represent a distinct historical conjuncture.

Most theorizing about the second question has used the conventional classification of types or forms of interorganizational relations that parses them into hierarchies, markets, and networks. This trichotomy has its roots in institutional theories in economies descending from Ronald Coase, William Ouchi, and Oliver Williamson. Nonetheless, there has been growing recognition of the fact that many arrangements or relationships that exhibit the formal features of markets or networks are, at the level of their informal institutions, more genuinely hierarchical. Conversely, it has long been recognized that many formally hierarchical partnerships are, in their informal institutions, highly individualistic. Finally, according to Perri 6, Nick Goodwin, Edward Peck, and Tim Freeman, it has been found that the concept of a "network" increasingly does not correspond to any single institutional form.

Authors have differed in the extent to which they have sought to use the trichotomy for truly theoretical and explanatory purposes or the extent to which it has been used mainly descriptively as a coding schema. For Koen Verhoest, Geert Bouckaert, and Guy Peters, the three institutional forms achieve their explanatory force by enabling the operation of what they describe as *mechanisms*, a term that in their usage largely corresponds to policy instruments used for control, incentive, or persuasion and for accountability. Some, such as Sullivan and Skelcher, propose positive feedback explanations, such that prior and latent institutional patterns are reinforced during periods of deliberate effort to cultivate greater coordination: This sits slightly uneasily with their negative feedback theory in answer to the first question, and reconciliation depends on very exact periodization that has not yet been tested with detailed historical research.

There have been many attempts to engage in theory development in the field using concepts of trust and networks. It is common to find authors, such as Robert Agranoff, claiming that coordination is best facilitated by nonhierarchical, high-trust consensual managerial practices, although it is rather rare to find close comparative analysis of cases where more hierarchical approaches are used, holding the service field and political jurisdiction constant; this leaves continuing question marks over these kinds of claims.

It is common to find small-*N* studies used to develop necessary-condition hypotheses about the extent of prior trust or prior operation of certain patterns of networks *of organizations* before we can expect to see even administrative implementation of programs being sustained, let alone positive impact on outcomes. Other studies, such as that by Jeanette Moore, Robert M. West, Justin Keen, Mary Godfrey, and Jean Townsend, or by H. Brinton Milward and Keith Provan, are concerned with networks of referral *of clients* between organizations, and they draw on sociometric approaches to network analysis to offer predictive understandings of these relations between organizations and services.

Eugene Bardach, presenting his argument relatively informally, has offered a strategy for answering the second question by reference to a thesis that assumes a rather strongly individualistic institutional character of organization, arguing that such a setting allows individual entrepreneurs to broker their way into fluid patterns of interorganizational relations, in which trust is taken to be a largely pragmatic affair. In a similar vein, there is a large body of work inferring from qualitative work on the role of particular individuals that "leadership" is a necessary condition for success, although quite what types of leadership have just which effects and in which circumstances remains ill specified.

The fourth question, about conditions for "success," is most frequently addressed—for example, by Bob Hudson, Brian Hardy, Melanie Henwood, and Gerald Wistow—using theories of organizational process, ranging from theories of the life cycles of initiatives through to more eclectic approaches that combine institutional constraints with models of interaction processes and organizational capability theory.

Opposing Principles

Advocates of coordination argue that the six concerns discussed above (in the "Goals and Means" section of this entry) are problems to be combated, as a matter of priority. But by no means do all governments or policymakers agree. For many, conflict, contradiction, and competition may be regarded as broadly healthy, and fragmentation and duplication might be regarded as losses that are acceptable as the price of other important gains. For, indeed, coordination is directly opposed to a number of principles for the organization of policy making and of public services that have long had, and continue to have, powerful advocates.

Four of the commonest principles that can and sometimes have implied contrary principles of organization might be called (1) *specialism in accountability*, (2) *specialism in creative tension*, (3) *liberal constitutionalism*, and (4) *populism*.

Specialism is the argument that organizing the principal accountabilities of policy making and public services around functions, by which we conventionally mean the administration of discrete services that call for particular bodies of knowledge, skill, professionalism, and so on, is the most effective way either to ensure efficient use of inputs or, perhaps, to secure the responsiveness of the services to the will of the government of the day. It is generally much more straightforward to attach estimates of costs and control of expenditure to inputs than it is to attach them to outputs or outcomes, though it is true that outcomes such as upward social mobility, health status, levels of educational attainment, the rate of unemployment, or the prevailing level of the fear of crime are each affected by the work of many agencies, departments, and services. But measuring the contribution of each to the outcome in a way that enables accountability or managerial control is extremely challenging; indeed, the particular relationships may not be very stable over time, yet systems of managerial control have to be kept in place for some time if we are to be able to tell whether they are capable of being implemented at all, let alone if we want to gather evidence about whether they might or might not be effective. We can, in short, manage and hold accountable each of the services for what they can do, but we will find it difficult to manage and hold each to account for what they contribute to what all should care about. Second, the specialist position argues that trying to manage and hold them to account for shared outcomes in practice results in the blurring of imperatives, scope for buck-passing, and the blunting of the commitment to the best use of specialist expertise, which is often critical to the best use of the professionals of medicine, policing, teaching, counseling, land use planning, building, and even administering of payments and processes that make up the interventions supplied by the public services. In the reforms of public management that were adopted in the central government in Britain and New Zealand and many other administrations during the 1980s, the argument of specialism was widely deployed, often under the banners of "managerial focus" and the "dedicated agency."

A second kind of argument for specialism directly contests the claim with which this entry began: that better outcomes result from reducing tensions between the policies implicitly or explicitly pursued by agencies. On the contrary, it could be suggested that it is precisely these tensions between the goals of different agencies that spur thought, innovation, adjustment, development, and sophistication related to both means and ends in those agencies. Tension, according to this argument, can be creative. At any one time, it may appear statically inefficient, because managing tensions must always take up managerial time and often cause no little irritation. It is easy to fall into the trap of imagining that if the tensions did not exist, then that time would be better deployed. However, this view would suggest that it is not necessarily or even typically true. Rather, too much pressure for coordination may undermine the creativity that, over the longer term, might produce much greater dynamic efficiency, which would more than compensate for any loss of static efficiency.

The liberal constitutionalist argument holds that both the coordinators and the specialists are misguided in their pursuit of either effectiveness or efficiency. Instead, the liberal constitutionalist fears the excess and abuse of governmental power above all and seeks to harness conflict, contradiction, competition, and incoherence as sources of checks and balances and as a mutually countervailing force for the competing imperatives. Based on this view, the real threat is not that government may not work very well but that it might become too powerful; we should be prepared to accept the consequences for any measures used to derail it, in order to prevent its becoming overweening.

The populist argument is that government should be organized around the principle that the people should be able immediately to understand its organization and that the units to which accountability, budgets, legal powers, and other resources are granted should be defined around things that the public can instantly recognize as having to do with tangible services. Organizing around grandiose and abstract outcomes and interdepartmental committees, creating crosscutting matrices of partnerships and joint budgets, and combining professions and services under headings that correspond to no single category of work may all be valuable instruments if all we ever cared about were that government produce results that can be measured statistically (see below on instruments of coordination). However, for the populist, the most important principle for the organization of government is not the achievement of results that are measured in that way but the intelligibility and immediacy of the organizational system of government to the governed and the ability of the governed to look into that system and readily appreciate what it does. If for the specialist, loss of managerial control over inputs for efficiency is the great risk, and if for the liberal constitutionalist, the overweening ambition and intrusion of government is the primary threat, then for the populist, the worst prospect is a system of government that is opaque and technocratic. In practice, populism has often led to an emphasis on organizing government according to the principles of clientele, territory, or jurisdiction or visible service function.

Readers of the history of the American Revolution will recognize the roots of specialism, liberal constitutionalism, and populism in at least some of the ideas associated with the names of Alexander Hamilton, James Madison, and Thomas Jefferson, respectively. The enduring political significance of these rival traditions, and the importance of media and pressure group activity to further their claims, forms part of the explanation for the fact that the periods when coordination is a central strategic commitment of reforming governments tend not to last very long. Sooner or later, however strongly they are committed to coordination, politicians must begin to give some weight to some of these other considerations. Bluntly, politics must care about many things in government other than the question of whether or not it works very well.

Instruments

The range of available policy instruments, mechanisms, tools, or administrative arrangements by which coordination can be pursued is very wide indeed. However, for the sake of convenience, they can be grouped as follows:

Bringing Together Inputs. This approach to the promotion of coordination spans the creation of budgets held jointly by two or more departments or agencies for a defined purpose: the creation of common information systems, programs of joint training, and individual posts that report to two or more agencies and pursue goals for which each agency has some part of the overall responsibility.

Bringing Together Organizations. Although one extreme of the distribution on this group of instruments includes the full merger of whole departments and agencies, there is a plethora of other ways in which organizations can be brought together. Joint matrix management is sometimes possible for organizations that have not fully merged but where each is asked to work toward goals previously expressed principally within another organization. The commonest instrument is the formal partnership, with its own joint board, but there are also several kinds of joint subsidiaries for joint ventures, all of which may be temporary or permanent and may cover all or only some of the functions of the participating agencies.

Bringing Together Accountability and Authority. Holding agencies to account for their contribution to achievements or problems for which each has some responsibility can be done in a variety of ways. Financially, the release of funds can be made conditional on the making of certain kinds of efforts and commitments. Managerially, executive oversight systems can be put in place that provide for review and accountability of teams and individual managers on the contribution of their work to goals that run beyond those of their immediate function. Parliamentary scrutiny can be exercised in a manner that does not confine itself to the oversight of its particular departments by each committee. Panels of service users can be recruited and supported to scrutinize the performance of many services to meet the needs or achieve the goals that they care and are informed about.

Types and Levels

Coordination takes its place in an extensive lexicon of cognate terms. At least some of the commoner ones are noted below. These terms can be grouped under the four headings shown below. The definitions of these cognate terms that are presented serve to display some of the main ways in which scholars have developed typologies and taxonomies for types of strategic approach, levels of coordination, and policy instruments used to pursue those different approaches.

Strategic Approaches

Horizontal government: a synonym for coordination

Joined-up government: coordination to combat inconsistency in activities and goals

Holistic government: coordination to pursue settlements in which potentially rival imperatives are made mutually reinforcing in pursuit of enhanced outcomes

Levels

Whole-of-government approach: coordination relying principally on bringing together accountabilities and oversight authority

Collaboration: any kind of coordination that involves bringing together either inputs or organizations

Joint working: collaboration on specific programs rather than for the whole pattern of relations between organizations

Integration: coordination that emphasizes bringing together organizations in especially tight ways—for example, through long-term, tightly supervised strategic partnerships or even mergers

Interdepartmental working: coordination at the tier of central government policy formulation work

Multiagency working: coordination in which the participating agencies retain their distinct identities

Interagency working: coordination that seeks to blur the distinct identities of the participating agencies

Interprofessional working: coordination at the level of the professions rather than the organizations

Instruments

Partnership: a formal structure with a constitution, a board, legal powers, a legal personality, and a specific task, in which participating organizations take full membership

Collaborative: a small team comprising professionals from several distinct professions working together with a distinct clientele or on a series of consecutive processes

One-stop shop: a "Main Street" business area–based or web-based facility in which citizens can access a wide range of services in one place

Problems Targeted

Crosscutting issue: any issue addressed by coordination that affects a wide range of different agencies, organizations, and professions

Wicked issue: any issue addressed by coordination that affects a wide range of different agencies, organizations, or professions; where the measurement of success in addressing the problem raises fundamental methodological problems; and where the causality by which interventions work appears to be nonlinear

Evidence of Effectiveness

There is a truly vast body of literature—including official governmental, practitioner, and peer-reviewed research texts—that claims to offer evidence for and against the effectiveness either of some of the particular recent programs or else of the uses of particular instruments. Unfortunately, much of this literature is rather unsatisfactory. Most studies are not conducted to evaluate effectiveness against any clear alternatives used for comparison or control. Indeed, it is not always clear just what the relevant practicable alternative program or instrument might be, against which comparison ought to be made. Before-and-after comparisons have well-known methodological weaknesses for the purpose of determining effectiveness. Moreover, most studies examine very small numbers of initiatives, limiting the scope for comparison. Because most of the outcomes for the environment, poverty, employment, health status, fear of crime, rates of reported crime, and so on are affected by a great many other factors than those of the deployment of instruments of coordination, for which few studies can realistically control, it is very difficult to place much confidence in the claims made. Recognizing these difficulties, most authors offer only process evaluations, rather than making claims for outcome evaluations. The process evaluation literature is replete with lists of "barriers" or "obstacles," either to the possibility of attempting particular programs of coordination or to the use of certain instruments at all or else to the chances of achieving satisfaction in the self-reports of the practitioners studied in the process. It is very difficult to know what to make of these many lists. For many of the entries have to do with things such as apathy, inertia, mistrust of others, protection of bureaucratic turf, risk aversion, concern for professional status, and fear of loss of control of budgets or other resources. Many of these things are either among the eternal verities of organizational life or else "barriers" to attempting anything and hardly specific to the instruments of coordination. At this stage in the development of the research, therefore, it is hard to be confident that much has been learned from the studies conducted in the 1990s and the 2000s that was not already known from the many, much earlier studies on the difficulties experienced in the joint implementation of policies in earlier decades—for example, from Jeffrey Pressman and Aaron Wildavsky's famous (1973/1984) study of the difficulties of multi- and interagency working in U.S. efforts to implement some of the "Great Society" economic development programs launched by the Lyndon Johnson administration in the 1960s.

Perri 6
Nottingham Trent University
Nottingham, United Kingdom

See also Policy Analysis; Policy Evaluation; Policy Formulation; Policy Framing; Policy Instruments; Policy Process, Models of

Further Readings

Agranoff, R. (2003). *Leveraging networks: A guide for public managers working across organizations.* Arlington, VA: IBM Endowment for the Business of Government.

Bardach, E. (1998). *Getting agencies to work together: The theory and practice of managerial craftsmanship.* Washington, DC: Brookings Institution Press.

Coase, R. H. (1937). The nature of the firm. *Economica, 16*(4), 386–405.

Hood, C. C. (1998). *The art of the state: Culture, rhetoric and public management.* Oxford, UK: Oxford University Press.

Hudson, B., Hardy, B., Henwood, M., & Wistow, G. (1999). In pursuit of inter-agency collaboration in the public sector: What is the contribution of theory and research? *Public Management, 1*(2), 235–260.

Milward, H. B., & Provan, K. G. (1998). Measuring network structure. *Public Administration, 76,* 387–407.

Moore, J., West, R. M., Keen, J., Godfrey, M., & Townsend, J. (2007). Networks and governance: The case of intermediate care. *Health and Social Care in the Community, 15*(2), 155–164.

Ouchi, W. G. (1980). Markets, bureaucracies, and clans. *Administrative Science Quarterly, 25*(1), 129–141.

Pollitt, C. (2003). Joined-up government: A survey. *Political Studies Review, 1,* 34–49.

Pressman, J., & Wildavsky, A. (1984). *Implementation* (3rd ed.). Berkeley: University of California Press. (Original work published 1973)

6, P., Goodwin, N., Peck, E. W., & Freeman, T. (2006). *Managing networks of twenty-first century organizations.* Basingstoke, UK: Palgrave Macmillan.

6, P., Leat, D., Seltzer, K., & Stoker, G., (2002). *Towards holistic governance: The new reform agenda.* Basingstoke, UK: Palgrave Macmillan.

Stinchcombe, A. L. (1990). *Information and organization.* Berkeley: University of California Press.

Sullivan, H., & Skelcher, C. (2002). *Working across boundaries: Collaboration in public services.* Basingstoke, UK: Palgrave Macmillan.

Verhoest, K., Bouckaert, G., & Peters, B. G. (2007). Janus-faced reorganization: Specialization and coordination in four OECD countries in the period 1980–2005. *International Review of Administrative Sciences, 3*(3), 325–348.

Williamson, O. E. (1986). *The economic institutions of capitalism: Firms, markets, relational contracting.* New York: Free Press.

Corporativism

Despite a centuries-long debate on corporativist forms of political representation and statehood, the concept of corporativism (corporatism) is still elusive, owing to insufficient definition and ideological controversy. This entry discusses the definition and the basic idea of corporativism with reference to its historical roots in the Middle Ages and its subsequent developments up to the 19th and 20th centuries. The more recent Catholic element and the secular one are also presented. Finally, the authoritarian features of corporatist rule, especially in Latin America but also in other parts of the world, are discussed.

The Notion

In its most basic meaning, corporativism refers to a political power structure and practice of consensus formation based on the functional representation of professional groups. Associations of farmers, craft workers, industrialists, laborers, lawyers, doctors, or the clergy (churches) act as self-governing bodies on their own behalf and as intermediaries between the government and their members. Political status and rights are attached to occupation and group membership and, thus, differ from those of modern citizenship and equal representation of individuals in parliaments related to a certain territory. The concept of corporativism dates back to the medieval estates and guild system. It found renewed attention among romanticist philosophers in the Germanic world during the 19th century as a remedy against social uprooting in the wake of industrialization and class conflict. The Catholic social teaching has drawn inspiration from corporativist ideas, as have a number of authoritarian regimes in Europe and South America. As a political ideology, corporativism has been fiercely rejected by liberals as a movement that would elevate collectivist corporate bodies to the cornerstones of politics and the economy while denying individual representation, civil liberties, free competition, and democracy. Socialists and Communists fought the idea as a particularly oppressive variant of capitalist class rule that would eliminate the Left and control the working-class masses with a carrot-and-stick approach. Not the least, it was the modern state itself and its ideal of universal citizenship, territorial instead of occupational representation, and indivisible sovereignty that ran counter to the establishment of corporate power holders acting as intermediaries between governments and segments of the society. Nevertheless, modern welfare states bear some features of corporativism and neo-corporatist policy making that have become manifest in networks and negotiations between state administrations and powerful corporate actors such as business associations or labor unions in fields such as industrial and social policy. In contrast to such current patterns of corporatist policy making, the term *corporativism* (corporatism, *corporativismo, corportativisme*, and *Korporativismus*) denotes a specific political philosophy and controversial ideology as well as a political regime type.

The Model and Idea of Corporativism

For centuries, debates on corporativism referred—directly or indirectly—to the medieval guild system. Being collective bodies (Latin *corpora*), medieval guilds served manifold economic, social, cultural, religious, and not the least political functions for their professional membership. Among them one finds the setting of standards for quality, prices and wages, education and work, caring for widows and orphans, representation on town councils, serving in courts and town militia, and maintaining charitable institutions such as hospitals, orphanages, poorhouses, and more. With the rise of the modern state and capitalist economies, those tasks became subject to royal law, state administrations, or markets. The decline of medieval corporativism began at the turn of the 14th century. It was driven not by absolutist claims to sovereignty at first but by recurrent pestilence, severe famine, and extreme weather as well as the spread of Renaissance humanism and Reformationist ideas about religion, science, and society that began to shake the old corporate order of the High Middle Ages and paved the way for a growing emphasis on individualism, centralized territorial rule, and secular supremacy.

Medieval Corporativism

Elements of corporativism appeared in many different sites of medieval political philosophy, though Johannes Althusius (1563–1638) was the first to formulate a comprehensive theory of a corporativist or, in his words, "consociationalist" constitution where the medieval order had already lost ground against new thoughts of monarchic absolutism and indivisible territorial sovereignty. Antony Black refers to Althusius as being one the "few great theorists of corporatism" (Black, 1984, p. 141), providing us "with perhaps the most substantial exposition of guild ideas ever known" (p. 131). The universal commonwealth (*consociatio universalis*), Althusius proclaims, has to be understood—in his own words—as "an association inclusive of all other associations (families, collegia [guilds], cities, and provinces) within a determinate large area and recognizing no superior to itself" (Althusius, 1603/1964, chap. 12). In conceiving the social contract as a real pact among corporate legal entities that compose society, Althusius differed from his near-contemporary Thomas Hobbes, who thought of a single agreement, entered into by individuals, who commit themselves to an absolute subjection to a common power: "one Man or one Assembly of men, that may reduce all their Wills, by plurality of voices, unto one Will" (*Leviathan*, Part II, chap. 17). Althusius had a notion of shared sovereignty that stands in deep contrast not only to Hobbes's unitarism but also to Jean Bodin's doctrine of monarchical sovereignty. Due to his emphasis on corporative autonomy, the subsidiarity principle, and the multilevel character of his constitutional system, Althusius is now reputed for being an early-modern protagonist and forerunner of federalism.

Corporativism has been criticized for its emphasis on collective, instead of individual, autonomy. While this is true for authoritarian concepts connected with various semicorporativist but in fact centralist regimes of the 19th and 20th centuries, early-modern "corporativist" thinkers like Althusius and Marsilus of Padua fought not only for a decentralized polity but also for popular sovereignty. In formulating that the sovereign rights of the people were inalienable, unassignable, and imprescriptible, Althusius rejected Bodin's notion of sovereignty according to which the people enter into the social contract with the monarch as a collectivity and not as free individuals. This is not surprising if one considers that medieval corporations, both in constitutional theory as in the everyday life of guilds and communes, was marked by a constant tension and interplay between the claims of individuals and those to be found on the level of corporations and their interrelations—a factually conflict-ridden power structure that fed unitarian anticorporativist political sentiments in the long run.

In serving the public and private needs not only of corporate status groups but also of communities at large, medieval corporativism ideologically aimed at an organic whole and, thus, was devoid of modern differentiations such as those between the collective and the individual, government (state) and society, politics and religion, or the public and the private sphere. The antimodernist overtone of corporativism derives from its organicist idea of segmenting society into an articulate multiplicity of interrelated semiautonomous corporations regulated by the principle that men should live freely only in the narrow sphere of their God-given social status.

Challenging the Liberal Paradigm

The medieval model of a static organic order received new attention from philosophers and state theorists during the 19th century. The dark side of rapid industrialization, together with the decay of traditional social security mechanisms and an emerging class conflict, resulted in intellectual attempts to find alternatives to the individualist market-liberal paradigm. Corporativist ideas came up again after the French Revolution, particularly regarding its failures and disappointments of widespread beliefs in a republican solution. They have to be seen as an attempt to reconstruct intermediary corporations as moral institutions to support communitarian politicization and individual orientation in times of rapid social change and as a barrier against social uprooting. Corporativist thoughts in the works of Johann Gottlieb Fichte, Adam Müller, and G. W. F. Hegel (see Tetsushi Harada, 1989) should help fill the gap between a society of individuals and the governmental administration with semiautonomous communal institutions that would eventually strengthen the state as an embodiment of the general interest. Social stabilization in the face of revolutionary threats has been the one—conservative—facet of such concepts, whereas the protection of craft workers, unskilled workers, and industrialists against social and commercial threats represents a more constructive if not humanistic concern. For Hegel (1770–1831), corporativism was the solution to the problem of an increasingly atomized society. Starting from the assumption that the rule of law is based on the need to articulate modern civil society as the realm of particularity, on one hand, and the state as the concrete form of moral generality, on the other, corporations are meant to embrace particularity from below and generality from above and, thus, to function as integrating links between civil society and the state. In his critique of Hegel's *Philosophy of Right*, Karl Marx identified the Achilles heel of Hegel's corporativist design exactly in the ambiguity of corporations (Estates) as intermediate power holders:

> While the Estates, according to Hegel, stand between the government in general on the one hand and the nation broken up into particulars (people and associations) on the other, the significance of their position . . . is that, in common with the organized executive, they are a middle term. Regarding the first position, the Estates represent the nation over against the executive, but the nation *en miniature*. This is their oppositional position. Regarding the second, they represent the executive over against the nation, but the amplified executive. This is their conservative position. They are themselves a part of the executive over against the people, but in such a way that they simultaneously have the significance of representing the people over against the executive. (Marx, 1978, p. 68)

Corporatist theories have always struggled with the dilemma of intermediate political corporations to represent and discipline the demands of their membership at the same time. One speaks of "societal" or "state" corporatism depending on whether a bottom-up representative approach or a top-down disciplinary one prevails.

During and after Hegel's time, a number of scholars—Otto von Gierke (1841–1921) and Émile Durkheim (1858–1917) in particular—suggested corporativist associations and intermediation as a remedy against the disorder, social anomie, and isolation resulting from a growing division of labor and the crisis of traditional institutions of solidarity.

Modern Corporativist Ideologies

Besides being a theoretical concept and regime type, corporativism has been perceived first and foremost as a political ideology, just like liberalism and socialism—the other great "ism," as Howard J. Wiarda called it. Among contemporary secular ideologies, it is not only the oldest but also the most manifold and inconsistent one. It is difficult to draw a clear dividing line between individualist liberalism and collectivist socialism since corporativist thoughts and methods took up elements of market liberalism as well as economic planning. It was a common trait of 20th-century corporativism to protect private property and simultaneously to fight against free competition in markets and politics. Experiences in Europe, Latin America, and Asia clearly show that authoritarian corporativism together with rightist populism has served as an instrument of conservative rule in periods of intensified class struggles and leftist revolts.

Papal Encyclicals

The Roman Catholic school of thought on the social, political, and economic order of capitalism, as put down first in the Papal encyclical *Rerum Novarum* (Of New Things, 1891) and supplemented by *Quadragesimo Anno* (In the 40th Year, 1931), above all builds on the corporativist principles of solidarity and subsidiarity. *Rerum Novarum* has to be seen as a response to the social instability and labor conflicts of its time. It advocates a "natural right" to form associations and engage in free collective bargaining and makes it the duty of governments to protect the poor. Faced with strong communist and fascist movements in Europe, *Quadragesimo Anno* also concentrates on the dangers to human freedom and dignity arising from unrestrained capitalism and totalitarian communism.

The Catholic plea for an organic and hierarchical reconstruction of industrial societies was inspired by, and in turn influenced by, the works of a number of 19th-century social philosophers from Italy (Frédéric Ozanam and Luigi Taparelli D'Azeglio) and France (Philip Buchez, Count Albert de Mun, and Count La Tour du Pin). They imagined

> Catholic Guilds where the interest of the profession is superior to private interest, where antagonism between capitalist and workingman gives way to patronage exercised in a Christian spirit and freely accepted. . . . It is always the same thought: limit competition, associate common interests, impose upon the employer the duty of patronage, uplift labor and the condition of the laborer. (Albert de Mun, 1847, as cited in Moon, 1921, p. 99)

Though full of normative ideology and never free of theoretical inconsistency, the social Catholic movement succeeded in maintaining a fairly coherent political reform program that is still focused on corporativist attributes of solidarity and subsidiarity. The insistent demand for subsidiary autonomy of the social sphere resulted in a somewhat muted critique of totalitarian fascism in *Quadragesimo Anno*. The Catholic Church, however, never did endorse—as German constitutional lawyers feared—a decomposition of statehood in the name of the subsidiarity principle.

Secular Corporativist Ideologies

Among secular corporativist ideologies, the idea of *guild socialism* found a number of supporters among intellectuals as well as workers and labor unions in Britain. Functional representation was one of the most distinctive doctrines of British guild socialism, according to which the population should be represented both as producers and as consumers in a multilevel system. Municipalities, regions, and the national state would be governed by two chambers: one elected by professional guilds and the other by territorial constituencies. George D. H. Cole, who formed the British National Guilds League in 1915, proposed a central guild congress that should be the supreme industrial body, standing for the people as producers in the same way as parliament will stand for the people as consumers. In advocating the cosovereignty of workmans' guilds and the representative government guild, socialism rejected the traditional notion of sovereign statehood. The concept had been indirectly inspired by Althusius's corporativist-cum-federalist medieval concept through the reception of Gierke's theory of associations. Part of Gierke's magnum opus *Das Deutsche Genossenschaftsrecht* (German Cooperative Law) was published in Britain in 1990 as *Political Theories of the Middle Ages* and subsequently gave an impetus to a newly emerging pluralist school of academic political thought, of which guild socialist ideologies attained the most far-reaching though rather short-lived political impact. The sudden decline of guild socialism after World War I had been attributed to a changing ideological climate that was no longer in favor of ideas requiring employers and business owners to share or give up control over industry.

The corporativist wave that began in the 1920s after World War I was conservative and ranged from right-wing to totalitarian political ideologies. Authoritarian nationalists in Franco's Spain, Salazar's Portugal, Dollfuss's Austria, Pilsudski's Poland, Vargas's Brazil, Perón's Argentina, or Calles's Mexico made use of corporativist theories to legitimize their claim to power as did totalitarian Fascists throughout Europe. During the 1930s, there were a number of fascist regimes not only in Mussolini's Italy and Nazi Germany but also in Hungary, Greece, and Romania as well as fascist movements and parties in almost all European countries. The corporativist and religious elements of fascism had been strongest in Italy and among (though also present within) the Falange movement in Spain and Romania. The attempt to violently infiltrate and control every sector of social life, together with their militaristic orientation, cast doubts over whether fascist dictatorships fall under the category of corporativism at all. Of course, there were a host of separate nongovernmental organizations controlled by government for all kinds of purposes, such as economic planning, leisure, sports, science, and education, but without any higher coordinating organs. Gaetano Salvemini, in his 1936 book *Under the Axe of Fascism*, noted that to search for real cooperation and genuine consultation taking place through corporatist institutions was like "looking in a dark room for a black cat which is not there." Contrary to the imagination produced by their propaganda machine, fascist regimes generally caused bureaucratic confusion and provoked the decline of constitutional and administrative order for the benefit of the revolutionary movement and violent oppression. Today, as Juan J. Linz has pointed out, there is widespread consensus not to subsume totalitarian fascism under the corporati(vi)st paradigm but rather to treat it as a separate regime type.

Authoritarian Corporativism

The term *authoritarian corporativism* originally referred to a variety of political regimes in Latin America. Most basically, it means that autocratic governments seek to impose a system of interest representation and intermediation on functional interest groups, especially labor unions, in order to deal with perpetual threats of industrial conflict and popular protest. Different structures and methods of incorporation across countries and over time, together with inconsistent or even lacking principles of political design, bear witness to the complexities involved in this political regime type.

Authoritarian corporativism emerged during the stages of late economic development from situations where weakly integrated societies, widespread clientelism, internal migration, and organizational fragmentation caused governments to reorganize, mobilize, and control socioeconomic groups in order to increase their predictability and eventually overcome economic stagnancy. Whereas the landed gentry, peasants, rural workers, clerics, and shopkeepers continued to be grouped territorially along vertical chains of patron–client relations—irrespective of class, race, caste, or region—up to the government level, new urban groups associated themselves along occupational and class lines on the national level. When, spurred by economic recession, the traditional landed, export, and commercial elite failed to check the manifold demands, political divisions, and modernization conflicts of such asynchronous societies "clientelism declined in favor of corporatism" (Guillermo O'Donnell, 1977, p. 67), and in many cases, it resulted in hybrid regimes of corporatist, populist, nationalist, and military authoritarian rule.

One can find similar sequences of events in many late-developing nations, with similar outcomes. South Korea, Taiwan, Singapore, Indonesia, the Philippines, China, and others have relied on various forms of functional representation though usually avoiding the label corporati(vi)sm because of its pejorative connotation of authoritarianism. Most East Asian developmental states borrowed heavily from Japan's earlier experiences.

Government-assigned farmer associations had been established there in the early 1900s for reasons of productivity improvement and political subordination. During the 1930s, the Japanese government reorganized the small- and medium-business sectors into peak associations, which were sanctioned to control their memberships through administrative decrees. Administrative guidance and cartelization have become symbols for the close state–society relations of the Japanese, but in the meantime, they have come to characterize a number of late-developing East Asian countries. In Taiwan, the government took over the farmers' associations that the Japanese colonial regime had established and afterward awarded an official quota of seats to functional associations in the country's—almost powerless—National Assembly. South Korea orchestrated its economic takeoff by means of government bureaus within a system of state corporatist arrangements, and China is about to top all previous attempts of corporativist control. Corporativist bodies dating back to the era of Mao Zedong have been revitalized as centers of guidance and legitimacy in economic and social matters. New associations in fields such as health, sports, culture, social welfare, or science have been established on the government's own initiative and—for the politically relevant—based on obligatory membership.

Apart from the more or less corporativist and authoritarian elements that can be located in current developmental states in Latin America, Asia, Southeastern Europe, or Africa, one can hardly find countries with a manifest corporativist constitutional background—with only a few exceptions. Among them is Croatia, with its long-established corporativist traditions that have led to a rather unique constitutional second chamber based on the corporative representation of trade unions, employers' and farmers' organizations, universities and colleges, craft workers, freelance professionals, and so on. Another case is Hong Kong, having up to half of the legislature elected by functional constituencies defined by professional occupations or economic sectors since 1985. Research on these cases suggests that strong corporativist regime elements as well as experiments with constitutional corporativism—other than neo-corporatist methods of policy coordination—have adverse affects on party development and the achievement of universal civil rights and have fostered legislative fragmentation through particularistic bargaining.

Roland Czada
University of Osnabruek
Osnabruek, Germany

See also Interest Groups; Neo-Corporatism; Pluralism; Representation

Further Readings

Althusius, J. (1964). *The politics of Johannes Althusius* (F. Carney, Trans.). Boston: Beacon Press. (Original work published 1603)

Black, A. (1984). *Guilds and civil society in European political thought: From the twelfth century to the present*. Ithaca, NY: Cornell University Press.

Harada, T. (1989). *Politische ökonomie des idealismus und der romantik: Korporatismus von Fichte, Müller und Hegel* [Political economy of idealism and romanticism: The corporatism of Fichte, Müller, and Hegel]. Berlin, Germany: Duncker & Humblot.

Linz, J. (2000). *Totalitarian and authoritarian regimes*. Boulder, CO: Lynne Rienner.

Malloy, J. M. (Ed.). (1977). *Authoritarianism and corporatism in Latin America*. Pittsburgh, PA: University of Pittsburgh Press.

Marx, K. (1970). *Critique of Hegel's philosophy of right* (J. O'Malley, Ed.). Cambridge, UK: Cambridge University Press. (Original work published 1844)

Moon, P. T. (1921). *The labor problem and the social Catholic movement in France: A study in the history of social politics*. New York: Macmillan.

O'Donnell, G. A. (1977). Corporatism and the question of the state. In J. M. Malloy (Ed.), *Authoritarianism and corporatism in Latin America* (pp. 47–87). Pittsburgh, PA: University of Pittsburgh Press.

Unger, J., & Chan, A. (1995). China, corporatism, and the East Asian Model. *Australian Journal of Chinese Affairs, 33,* 29–53.

Wiarda, H. J. (1997). *Corporatism and comparative politics: The other great "ism"* (Comparative Politics Series). Armonk, NY: M. E. Sharpe.

CORRELATION

Correlation is a statistical measure of the association between two or more variables. Two or more variables are associated if they change together (covary). In statistics, such simultaneous movement is called dependence. Thus, correlation is simply an indication of a lack of independence between variables. Measures of correlation can also indicate the direction of the relationship between variables—that is, whether it is positive or negative. In addition, while correlation alone is insufficient to establish a causal relationship, it often illuminates the potential for such. Below, the basic properties and applications of current correlation measures are discussed.

The modern statistical concept of correlation was initially put forward by Sir Francis Galton in a series of papers on heredity in the late 1800s. Correlation was subsequently refined in various works by Karl Pearson and G. Udny Yule. These initial works culminated in a statistic now called the Pearson product–moment correlation coefficient, Pearson's r. Pearson's r is perhaps the most widely used of all statistics because of its value as an indicator of correlation and its relationship to multivariate analysis, in particular linear regression. However, it is only one of several indicators of statistical correlation. Today, the concept of correlation is a broad term that comprises a host of indices used to measure both parametric and nonparametric association. Of the latter, Kendall's τ and τ_b, Cramér's V, and Goodman and Kruskal's γ are especially prevalent in applied statistics.

Pearson's r indicates both the degree and direction of linear dependence between variables. Direction in a linear relationship can take one of two forms: positive or negative. For example, take two variables X and Y. If high scores on X correspond to low scores on Y, there is a negative relationship. If high scores on X correspond to high scores on Y, there is a positive relationship. Similarly, degree measures the strength of a relationship between variables. Such a relationship might be strong or weak, where a strong correlation means that large values of X are associated with large values of Y and vice versa.

Correlation may be visualized via scatterplots—that is, by arranging the values of one variable on one axis and the corresponding values of another variable on another axis. In scatterplots, points falling close to a straight line imply strong correlation, whereas a cloud of points are more suggestive of a weak correlation or a lack of one. Similarly, if most points are in the bottom-left and top-right quadrants, a positive relationship is plausible, whereas if most points are in the top left and bottom right, a negative relationship is likely. A scatterplot may also detect nonlinear relationships between variables.

The Pearson r can be intuitively derived from a standardization of the covariance between variables. The covariance between X and Y is

$$\frac{\sum(X-\bar{X})(Y-\bar{Y})}{N},$$

where N is the sample size, and $\bar{X}$ and $\bar{Y}$ are the means of the variables. Covariance combines the deviations computed using the scores from each variable and accounts for the sample size in the denominator.

However, covariance does not control for the potential for substantial differences in the amount of deviation within each variable. Thus, to derive Pearson's r, the covariance is adjusted by the deviation of each variable. Simply divide the covariance by the product of standard deviations as given below:

$$\frac{\frac{\sum(X-\bar{X})(Y-\bar{Y})}{N}}{\sqrt{\frac{\sum(X-X)^2}{N}}\sqrt{\frac{\sum(Y-\bar{Y})^2}{N}}},$$

which, through some basic algebra, becomes the typical formula for Pearson's r:

$$r = \frac{\sum(X-\bar{Y})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{Y})^2(Y-\bar{Y})^2}}.$$

Pearson's r has the attractive property of lying between -1 and 1. Positive values signify positive relationships and negative values negative ones. The higher the absolute value of a measure, the stronger the relationship. Thus, a value of zero indicates that the variables are linearly independent of each other. A correlation of zero, however, does not indicate that the variables are necessarily completely independent. The relationship must be linear for Pearson's r to be an appropriate indicator of correlation. Curvilinear or alternative relationship forms will not be properly detected with Pearson's r. In addition, caution should be used when applying Pearson's r to small sample sizes.

Consider Figure 1 (next page) that presents scatterplots of a strong positive and a weak negative correlation between two pairs of variables for 23 countries that belong to the Organisation for Economic Co-operation and Development (OECD). In the left scatterplot, the tight grouping of countries' (X, Y) coordinates along the best fitting line suggests a strong positive correlation between countries' economic openness and growth in real gross domestic product (GDP). Pearson's r is .73, which confirms the relationship illustrated in the scatterplot. In the right scatterplot, the looser cloud of points suggests a weaker negative correlation between the percentage of women in parliament and the percentage of right-wing parties in cabinet posts. In this case, Pearson's r is −.51.

Pearson's r can be used to make statistical inferences about the population correlation ρ. Inference typically assumes a bivariate normal distribution, such that the relationship between X and Y must be linear, and X and Y must be continuous. Given a Pearson's r from a reasonably large sample, the null hypothesis of $\rho = 0$ can be tested with a t distribution, such that

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}},$$

where $n - 2$ represents the degrees of freedom.

The formula for Pearson's r is both simple and versatile, so much so that it can be expanded to relationships between variables at various levels of measurement. As a particularly useful case, consider the relationship between two ranked variables, which in the case of ties between variables' scores can also be computed with Pearson's r. When there are no ties, the formula can be simplified to the Spearman rank-order correlation coefficient:

$$r_{\text{rank}} = 1 - \frac{6\sum d^2}{n(n^2-1)},$$

where $d = X - Y$ or the difference in the ranked scores. Testing ρ for ranked variables is also performed via sampling from a t distribution with

$$t = \frac{r_{\text{rank}}}{\sqrt{(1 - r_{\text{rank}}^2)/(n-2)}}$$

and degrees of freedom $n - 2$.

Correlation indices can detect a relationship between variables even when no causal link exists between them. There are hosts of examples where two variables are perceived to correlate, because a third variable is causing them both to change together. This is called a spurious relationship. Thus, there is no direct causal relationship between the two variables, but there is a common dependence on the third. For example, when considering the correlation of partisanship with general attitudes on violence, the potential for a spurious relationship might be furthered by examining the role of gender, where women are more likely to join a left-leaning party and also oppose violence.

Given concerns about a spurious relationship, applied researchers have taken recourse to statistical control with partial correlation. Partial correlation measures the relationship between two variables while controlling for one or more variables. For example, the correlation of X with Y controlling for a series of control variables, $Z_i = \{Z_1, Z_2, \ldots, Z_k\}$, can be estimated as the correlation between the set of residuals, R_X and R_Y, calculated from the least squares regression of X on Z_i and of Y on Z_i, respectively.

Assuming that a single Z is the only variable that affects X and Y, the partial correlation coefficient, $r_{xy.z}$, can be simplified to

$$r_{xy.z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{1-r_{xz}^2}\sqrt{1-r_{yz}^2}}.$$

A hypothesis test of the population estimate, $\rho_{xy.z}$, can be performed with a t test, where

$$t = \frac{r_{xy.z}}{\sqrt{1-r_{xy.z}^2}}\sqrt{n-3},$$

where $\sqrt{n-3}$ is the degrees of freedom.

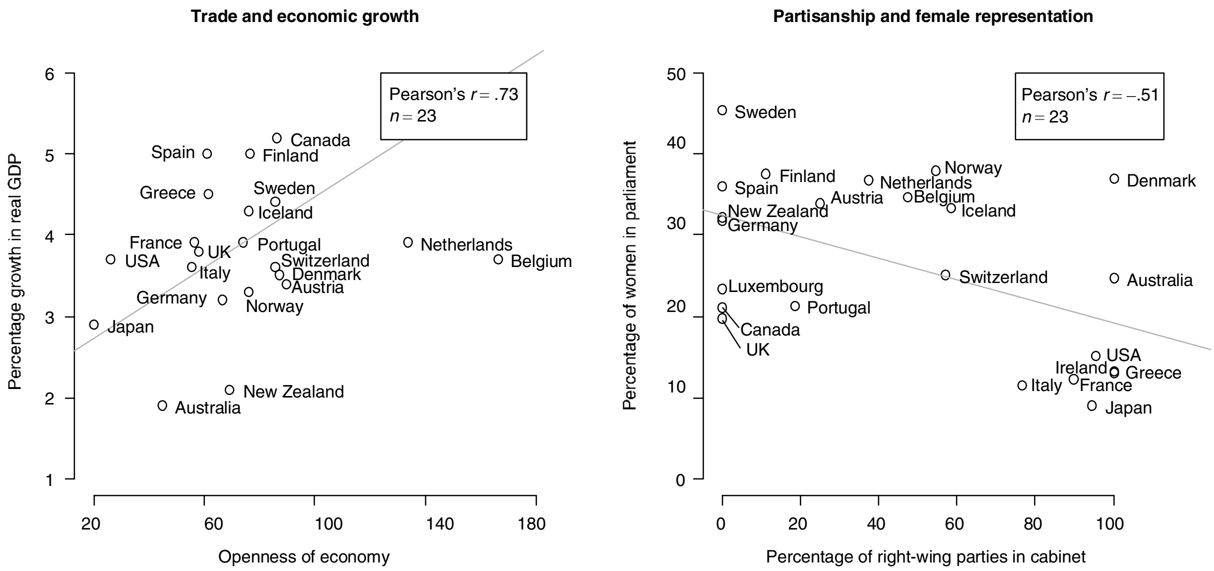


Figure 1 Scatterplots of a Strong Positive Correlation and a Weak Negative Correlation

Source: Data taken from Armingeon, K., Engler, S., Gerber, M. Leimgruber, P., & Beyeler, M. (2009). *Comparative political data set 1960–2007*. Bern, Switzerland: Institute of Political Science, University of Bern.

It may already be clear that correlation is closely related to multivariate analysis with least squares regression. Pearson's r essentially measures the direction and degree to which observations fall close to a line, which indicates a relationship between two variables. The best fitting line in such a relationship is the least squares regression line.

The regression line is arrived at by minimizing the squared errors with the function $\hat{Y} = a + bX + e$, where a and b are constants and e is the residual. Thus, in the regression context, Pearson's r denotes the potential for a stronger fit in the regression line. With a strong relationship, the sum of residuals is smaller, and the independent variable, X, explains more of the variance in the dependent variable, Y. In fact, squaring r, we get the coefficient of determination, r^2, in a regression equation. Therefore, r^2 is an estimate of the amount of variance in Y explained by the predicted relationship $a + bX$. Accordingly, the coefficient of nondetermination, $1 - r^2$, indicates the amount of unexplained variance in the model.

To make the connection clear, consider an alternative derivation of r^2 from a simple linear regression; r^2 is a ratio of the expected sum of squares to the total sum of squares, or

$$\frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y - \bar{Y})^2}.$$

This can be calculated from the residuals instead of the expected squares such that one less the residual sum of squares over the total sum of squares is the coefficient of determination:

$$r^2 = 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \bar{Y})^2}.$$

Because r^2 indicates the ratio of explained to total variance, it can be considered an estimate of the fit of the regression model. In a linear regression model, the coefficient of determination can range from 0 to 1. Returning to the example above, on the correlation between trade openness and GDP (see

Figure 1), the coefficient of determination would be simply $r^2 = .73^2 = .53$. In a regression context, this suggests that changes in trade openness explain about 53% of the variance in GDP growth. Likewise, according to the coefficient of nondetermination, about 47% of the variance in GDP growth is unexplained by this simple bivariate relationship.

Measures of correlation may also be nonparametric. Nonparametric correlation applies to relationships where the distributional assumptions above (e.g., drawing from a bivariate normal distribution) are not met. Oftentimes, this is the result of dealing with noncontinuous variables—that is, ordinal, nominal, or a mix of measurement levels. Similar to Spearman's rank-order coefficient above, Kendall's τ and τ_b are performed to test the correlation between ordinal variables. Kendall's τ counts all pairs of observations between two variables that are consistently above or below the rankings of correspondent pairs as concordant, or C. More formally, a C is counted every time the product of $(X_i - X_j)(Y_i - Y_j)$ is positive. Similarly, all pairs that are inconsistently above or below other pairs are counted as discordant and D. Kendall's τ is calculated to expose the difference between discordant and concordant pairs:

$$\tau = \frac{2(C - D)}{n(n - 1)}.$$

Here, τ ranges from -1 to 1, and the direction of the relationship is evident from the formula $(C - D)$, such that negative outcomes suggest a negative relationship and vice versa. For sample sizes greater than 10, a test of $\tau = 0$ can be done with a z test:

$$z = \frac{\tau}{\frac{\sqrt{2(2n + 5)}}{3\sqrt{n(n - 1)}}}.$$

When there are tied ranks in the data, slight derivations to the τ formula given above can account for ties. Kendall's τ_b and Stuart's τ_c require only a few additional terms to model the ranked ties. In the former, ties among the X values are collected in u, which enters the equation as

$$U = \frac{1}{2}\sum u(u - 1),$$

while ties in Y are collected in v, which similarly enters the equation as

$$V = \frac{1}{2}\sum v(v - 1).$$

In the latter, m is the smaller number of ranked categories in X and Y. Kendall's τ_b is then a calculation of discrepancy in concordant and discordant pairs while controlling for the number of ties in each variable:

$$\tau_b = \frac{C - D}{\sqrt{\frac{n(n-1)}{2} - U}\sqrt{\frac{n(n-1)}{2} - V}}.$$

Similarly, Stuart's τ_c takes into account tied ranks via the smaller number of ranked categories, m:

$$\tau_c = \frac{2m(C - D)}{n^2(m - 1)}.$$

Note that when there are many ties, both of the above measures will slightly underestimate the association. In that case, Goodman and Kruskal's gamma does a better job of correcting for several ties:

$$\gamma = \frac{(C - D)}{(C + D)}.$$

When there are no ties, this equation is equivalent to that for Kendall's τ.

The relationship between multichotomous variables can also be explored with correlation coefficients. Despite the prevalence of the contingency coefficient, C, Cramér's V is a more practical choice for dealing with this kind of data. Consider a three-by-three contingency table of two multichotomous variables: partisanship (Republican, Independent, and Democrat) and campaign advertisement exposure (television, Internet, and radio). To test the independence assumption between partisanship and campaign media, one applies the χ^2 test. However, the χ^2 test only provides information about the existence of a relationship, not the strength of a relationship. To acquire the degree of association between the variables and control for the ratio of rows to columns, one applies Cramér's *V*:

$$V = \sqrt{\frac{\chi^2}{n[\min(r, c) - 1]}},$$

where $[\min(r, c)]$ is the smaller of the number of rows or columns in the contingency table, which provides an accurate measure of association for multichotomous variables.

The correlation statistics presented here are only a sample of those most commonly used in social science. Because correlation plays a central role in understanding causal and spurious relationships, it has garnered a great deal of attention in the 100-plus years since its inception. In modern applied statistics, it remains one of the foundational concepts and the starting point for most practical research.

Dino P. Christenson
Ohio State University
Columbus, Ohio, United States

See also Quantitative Methods, Basic Assumptions; Regression; Statistics: Overview

Further Readings

Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statistical Science, 10*(4), 364–376.

Chen, P. Y., & Popovich, P. M. (2002). *Correlation: Parametric and nonparametric measures.* Thousand Oaks, CA: Sage.

Gibbons, J. D. (1993). *Nonparametric measures of association.* Newbury Park, CA: Sage.

Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis.* Englewood Cliffs, NJ: Prentice Hall.

Rodgers, J. L., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician, 42*(1), 59–66.

Weisberg, H. F. (1974). Models of statistical relationship. *American Political Science Review*, *68*(4), 1638–1655.

CORRESPONDENCE ANALYSIS

Correspondence analysis is a method for interpreting tabular data visually in the form of spatial maps in which the rows and columns of the table are depicted as points. Two forms of the method are common: simple-correspondence analysis (CA) and multiple-correspondence analysis (MCA). Cross-tabulations and raw categorical data in the social sciences are prime examples for being visualized by CA. Typically, a cross-tabulation is subjected to a test for association between the row and column variables using a chi-square test, for example. By contrast, CA visualizes the actual structure of this association, whether it is statistically significant or not. MCA generalizes this method to many variables, usually questions in a survey, showing how the response categories interrelate. In this entry, each approach is explained using data from an international survey on the role of government.

The theory underlying CA has its origins, as early as the 1940s, in the scaling of categorical variables (i.e., assigning numerical scores to their category levels) to achieve an objective such as maximizing their pairwise correlations or maximizing between-row or between-column variances. The geometric approach along with its name is derived from the French term *analyse des correspondances*, developed and popularized by Jean-Paul Benzécri and colleagues starting in the early 1960s. It is this approach that is presented here.

Simple Correspondence Analysis

As a first example, consider data from the International Social Survey Programme's Role of Government survey in 2006, in particular a question concerning government's contribution to the public health system. Respondents were asked whether government should pay "much more," "more," "same as now," "less," or "much less" for health services, and the responses were cross-tabulated with the respondents' interest in politics, in five levels (from 1 = *very much interested* to 5 = *not at all interested*). Table 1 shows these counts for three different samples, from France and the former West and East Germany, the latter still being kept separated for research purposes. In this table, the response categories "less" and "much less" have been combined because of very low frequencies of the latter response to the question. To simplify the present analysis, respondents with missing values have been removed. Missing data such as "can't choose," "don't know," or "refused to answer" can be coded as an additional level, either combined or separately, but these nonsubstantive responses often dominate the results, obscuring the interpretation of the substantive responses (Michael Greenacre & Jorg Blasius, 2006).

The first thing to note about CA is that it depicts the relative frequencies of response, called *profiles*. For example, the profile of French respondents with a high interest in politics (row **F1** of Table 1) is [73/229, 73/229, 65/229, 18/229] = [0.319, 0.319, 0.284, 0.079]—hence, a profile is a set of proportions summing to 1. The profiles of the French sample as a whole and the two German samples, as well as the respondents all together, called *average*

Table 1 Government Spending on Health

Country/		*Much More*	*More*	*Same*	*Less/Much Less*	
Political Interest		*He1*	*He2*	*He3*	*He45*	*Sums*
	F1	73	73	65	18	229
	F2	143	237	199	53	632
France	**F3**	111	204	171	37	523
	F4	45	88	95	20	248
	F5	19	27	22	1	69
	DW1	12	38	33	14	97
	DW2	30	88	83	20	221
West Germany	**DW3**	82	207	159	30	478
	DW4	37	99	44	15	195
	DW5	17	34	22	1	74
	DE1	9	16	13	3	41
	DE2	22	48	28	0	98
East Germany	**DE3**	65	112	48	3	228
	DE4	26	68	16	4	114
	DE5	8	23	5	2	38
	F	391	629	552	129	1701
		0.230	***0.370***	***0.325***	***0.076***	
	DW	178	466	341	80	1065
		0.167	***0.438***	***0.320***	***0.075***	
	DE	130	267	110	12	519
		0.250	***0.514***	***0.212***	***0.023***	
	ALL	699	1362	1003	221	3285
		0.213	***0.415***	***0.305***	***0.067***	

Source: Created by the author Michael Greenacre using data from International Social Survey Programme, Role of Government IV (2006). Distributor: GESIS Cologne Germany, ZA No. 4700.

Notes: Frequencies of responses to the question "How much should government pay for health services?" are cross-tabulated by political interest (from 1 = *very much interested* to 5 = *not at all interested*); samples are from France (N = 1701), the former West Germany (N = 1065), and the former East Germany (N = 519). Country totals and overall total, with corresponding relative frequencies, are given in the last rows. **He** = public health; **F** = French; **DW** = former West Germany; **DE** = former East Germany.

profiles, are given in the last rows of Table 1 in bold italics, and the data for group **F1** can be compared with these averages. Thus, compared with all French, those with a high interest in politics are proportionally higher in number in saying that government should spend much more on health (0.319 compared with 0.230); and they are proportionally fewer in saying that they should spend the same as now (0.284 compared with 0.325). This pattern is similar when compared with the total sample.

All the comparisons of the profiles with one another, and by implication with the averages as well, can be depicted optimally in one CA map, shown in Figure 1. A distance function, called the chi-square distance, is defined between profiles, and an approximate map of these distances is drawn. The chi-square distance recognizes that differences between high relative frequencies are inherently greater than those between low ones and adjusts for this disparity using the average of each response category (given in the last row of Table 1). For example, the chi-square distance between the first two rows, **F1** and **F2**, of Table 1, with profiles [0.319, 0.319, 0.284, 0.079] and [0.226, 0.375, 0.315, 0.084] is

$$\sqrt{\frac{(0.319-0.226)^2}{0.213}+\frac{(0.319-0.375)^2}{0.415}+\frac{(0.284-0.315)^2}{0.305}+\frac{(0.079-0.084)^2}{0.067}}=0.227.$$

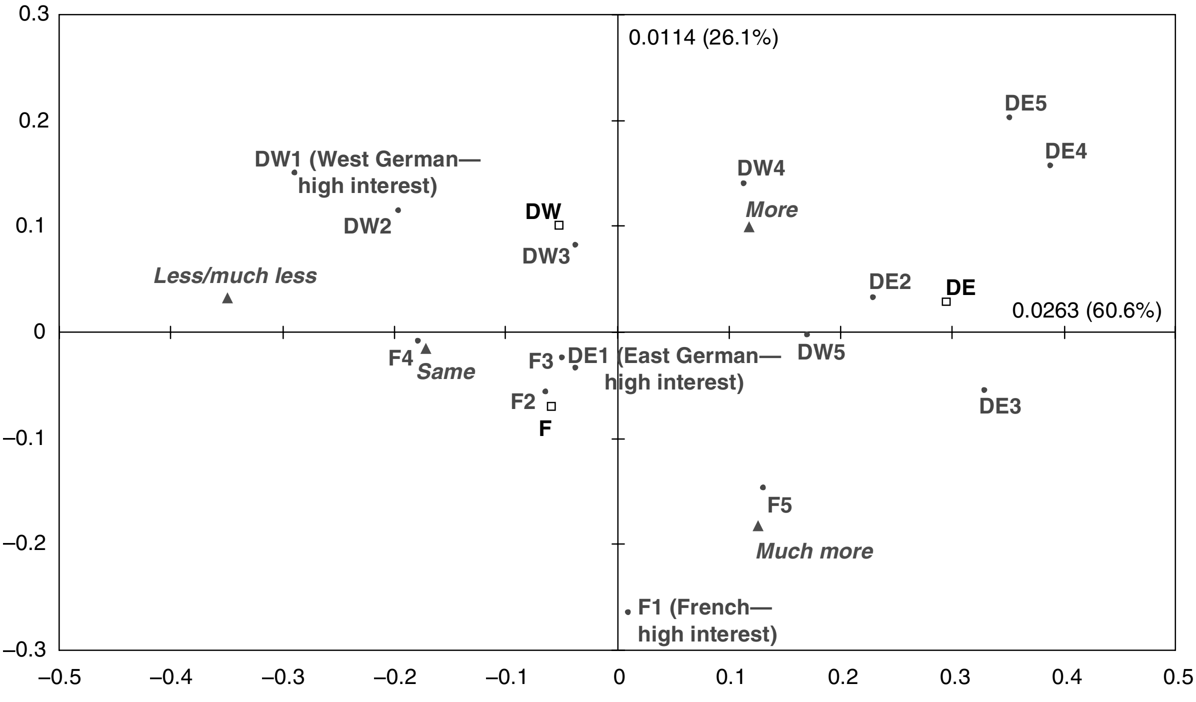


Figure 1 Correspondence Analysis Map of the Cross-Tabulation in Table 1.

Source: Created by the author Michael Greenacre using data from International Social Survey Programme, Role of Government IV (2006). Distributor: GESIS Cologne Germany, ZA No. 4700.

Notes: The average profiles for **F** = France, **DW** = former West Germany, and **DE** = former East Germany are shown as supplementary points. The overall profile (last row of Table 1) is the center of the map.

The smaller squared difference $(0.079 - 0.084)^2$ between the fourth category is increased by dividing it by the smaller average value of 0.067, compared with the larger squared differences in the other categories, which are divided by their respectively larger averages. It is this distance that the map of Figure 1 attempts to emulate in the positioning of the points **F1** and **F2**. The points **F**, **DW**, and **DE** in Figure 1, depicting the average profiles in each country, are called *supplementary points*: They are not included in the CA when computing the map but are added to the map afterward. The point **F**, for example, is the weighted average of the five points **F1**, **F2**, . . ., **F5**, where these five political interest points are weighted proportionally to their respective frequencies in the French sample.

Apart from the row profiles and their averages, the response categories of the spending question (column points) are visualized in the same map, and this provides the basic interpretation of the different regions of the map. Hence, the more to the right of the map, the more frequent is the attitude toward more spending, with the upper quadrant corresponding to "more" and the lower quadrant to "much more," while the more to the left of the map, the more frequent the attitude is toward spending the "same" or "less/much less." The center of the map corresponds to the average over all the samples studied, so by "more frequent" we really mean "more frequent compared with the average." Individual profile points can be compared: For example, the point **F1** corresponding to the group described above is the one that is the most in the direction of "much more" spending—it can be verified that there is no other row in the table that has a higher proportion of respondents having this opinion. The map allows an easy comparison of each of the five levels of political interest in each country as well as an overall comparison between the countries. East German opinion is well to the side of spending "more" and "much more," as opposed to West German opinion on the other side, along with the French. The French average, however, is lower down—all five French points are lower negative on the vertical dimension on the side of "much more." Of the five levels of political interest, the high level shows the greatest differences between the countries, with **F1** the most toward "much more," **DW1** the most toward "less/much less," and **DE1** in between. To illustrate further the style of interpretation of these maps, note that the East German groups **DE4** and **DE5** (low levels of political interest) must have higher proportions of opinions toward "more" spending than the West German group **DW4**, for example. It is not the proximity of the groups to the response category that indicates a higher propensity toward that category but rather how far out from the average (the center) the points are in the direction indicated by the response category—"more" in this case.

CA aims to visualize the relationships between the rows and columns of a cross-tabulation. A separate issue is to test whether this association is statistically significant or not. Chi-square statistics, for example, can be computed for the three 5×3 tables in Table 1 and are equal to 22.5 ($p = .033$), 28.0 ($p = .006$), and 24.1 ($p = .020$), respectively. For the entire 15×3 table, chi-square is equal to 142.9 ($p < .0001$); so from a statistical viewpoint, we are looking at significant differences both within each country and among them. In CA, the chi-square statistic divided by the grand total of the table is called the *inertia*, a dimensionless quantity that measures the total variation in the data: In this case, the inertia is equal to $142.9/3285 = 0.0435$. This inertia is decomposed along the axes of the CA solution in a similar way to the variance decomposition in principal component analysis. These parts of inertia are indicated on the axes of Figure 1, as also their percentages of the total inertia: 0.0263 (60.6%) and 0.0114 (26.1%), so that 86.7% of the inertia is accounted for, while 13.3% is not shown in the two-dimensional solution.

In summary, correspondence analysis of a cross-tabulation maps the association (measured by the inertia) between the rows and the columns of a cross-tabulation into a low-dimensional solution, facilitating the interpretation of this association. The method maximizes the inertia represented in the chosen solution, which is usually two-dimensional.

Multiple Correspondence Analysis

The simple form of CA described above is applicable to a single cross-tabulation, visualizing the relationship between two categorical variables. In MCA, there are typically several variables describing a particular phenomenon of interest, and the method's objective is to visualize how all these variables interrelate among each other. For example, consider not one but several questions about government spending: in addition to spending on public health (abbreviated as He), spending on environmental protection (En), law enforcement (La), education (Ed), defense (De), retirement pensions (Re), unemployment benefits (Un), and culture (Cu). There are eight questions in all, and each has the same five response categories (here we can keep the categories "less" and "much less" separated, because for some of the questions, there are quite high frequencies of the response "much less"). After the respondents with missing values on any of the questions have been removed, we have a total sample of 2,913 respondents: 1,486 French, 957 West Germans, and 470 East Germans.

There are several equivalent ways to define MCA: Here, we describe the method as an analysis of *dummy variables* for all the response categories. The original 2913 × 8 matrix of responses can be converted into a 2913 × 40 *indicator matrix* of dummy variables, where there are five columns per question for its five response categories. The responses are coded as 0s and 1s, with a 1 indicating the response category: For example, the response "more," which is the second response category, is coded as 0 1 0 0 0. MCA is the application of CA to this indicator matrix, leading to points for all categories of response and for all respondents, as shown in separate maps (A and B) in Figure 2. The method then visualizes the associations between these categories across the respondents: As in simple CA, categories that co-occur much more frequently than one would expect from their marginal frequencies will be grouped together, while those that have fewer than expected respondents in common will be placed far apart.

Correspondingly, respondents with similar response patterns will be located close to one another, and ones with different patterns will be located far apart. The substantive questions are again used to interpret the space—thus, it can be seen that the more the respondents (or groups of

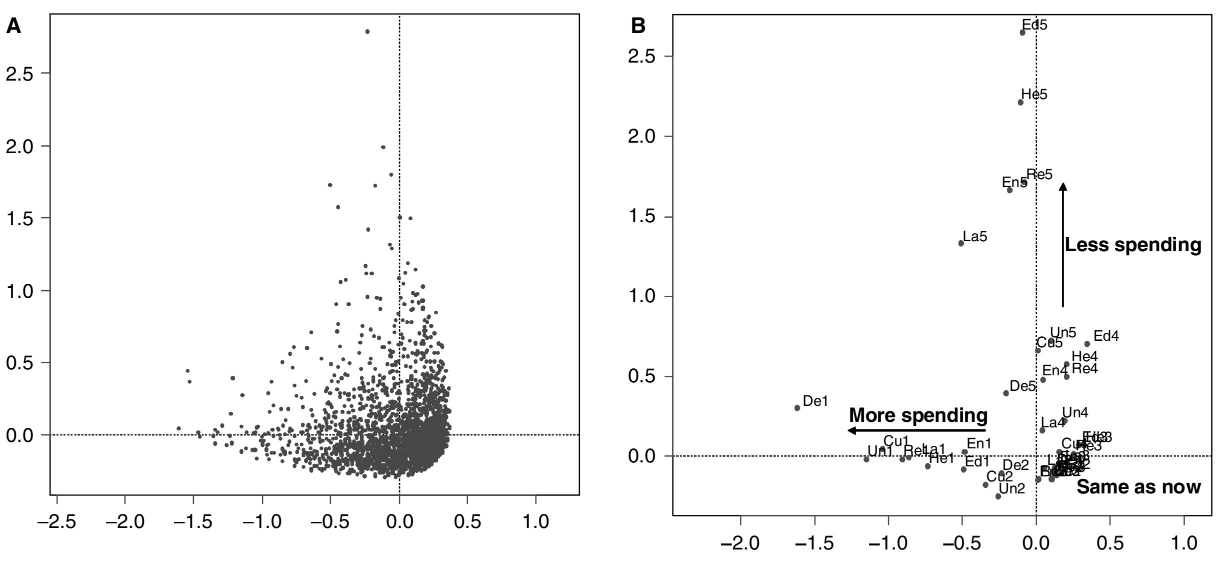


Figure 2 MCA of Responses to Eight Questions on Government Spending

Source: Created by the author Michael Greenacre using data from International Social Survey Programme (ISSP), Role of Government IV (2006). Distributor: GESIS Cologne Germany, ZA No. 4700.

Notes: Responses were obtained from 2,913 respondents from France and Germany (ISSP, 2006). Map A shows the respondents, and Map B shows the corresponding map of the response categories. **MCA** = multiple-correspondence analysis; **He** = environmental protection; **En** = environmental protection; **La** = law enforcement; **Ed** = education; **De** = defense; **Re** = retirement pensions; **Un** = unemployment benefits; **Cu** = culture.

respondents) are located up on the vertical axis, the less they advocate government spending, and the more they are to the left on the horizontal axis, the more they advocate spending. The most extreme views in these two directions are "spending much less on education" (point **Ed5**) and "spending much more on defense" (**De1**). The dense cloud of respondent points at the lower right, however, shows that the majority of responses are at or near the middle-response categories "same as now." If respondents tend in the opposite direction toward the upper left, this would reflect a type of polarization in their response pattern, combining the two attitudes captured by the two axes. For example, the position of **La5**, which is pulled somewhat into this quadrant, away from the vertical spread of the other "spend much less" categories toward the "spend much more" categories, implies that there are several respondents who—apart from advocating much less spending on law enforcement—are in favor of more spending on several of the other aspects. Another way of thinking about the correspondence between response categories and the respondents is that the position of each response category is the average of the respondents who gave that particular response. The relative positions of the categories thus reflect associations in the responses: For example, the category **Ed1** (spend much more on education) is much closer to the bunch of middle responses than **Ed5** (spend much less on education), showing that spending much more on education is not an extreme view but associated with an attitude generally associated with middle responses, while spending much less on education is an extreme attitude.

In the same way, a supplementary categorical variable, such as the country or level of political interest, can be added to the map afterward to see if it has any connection with the configuration of response category associations. Figure 3 shows the positions of the countries, genders, levels of political interest, and education levels as mean positions of the respondents in the corresponding category. In other words, the point **m** shows the average of the 1,538 male respondent points in Figure 2 and **f** the average of the 1,375 female respondents—so males are on average more in favor of spending less, and females on average are more in favor of spending more. Interestingly, the highest and lowest categories of political interest fall close to each other, together with the lower categories of education, all on the left side of the map toward more spending.

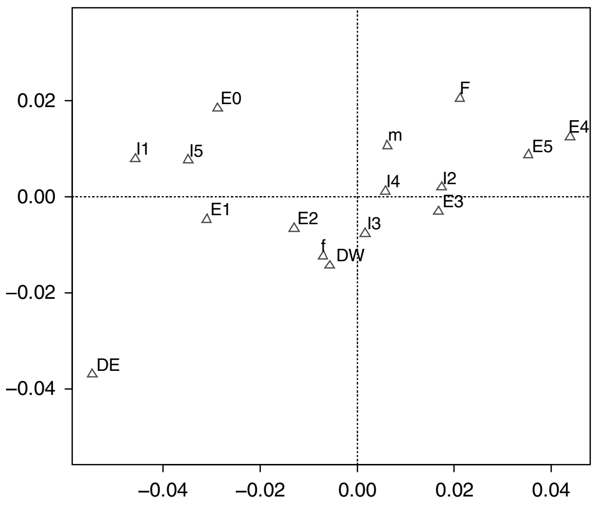


Figure 3 Average Positions of Groups of Respondents in Correspondence Analysis Map of Figure 2

Source: Created by author Michael Greenacre using data from International Social Survey Programme, Role of Government IV (2006). Distributor: GESIS Cologne Germany, ZA No. 4700..

Notes: **F** = France, **DW** = former West Germany, **DE** = former East Germany; **m** = male, **f** = female; **I1**, **I2**, . . . , **I5** = interest in politics, from high to low; **E0**, **E1**, . . . , **E5** = education level, from low to high.

Comparing the scale of Figure 3 with that of Figure 2, one can see that these differences are, in fact, quite small and possibly need to be qualified by some type of hypothesis test if the statistical significance of the findings is sought.

In summary, MCA analyzes respondent-level data on several categorical variables simultaneously in one spatial map. Although individual respondent points exist, it is the response category points that are of main interest. These category points are positioned to reflect optimally their associations with one another across the respondents. Additional categorical variables that classify the respondents into groups can be displayed as supplementary points after the map is constructed, to assist in the interpretation—these are the average positions of the respondents in the respective categories.

Michael Greenacre
Universitat Pompeu Fabra
Barcelona, Spain

See also Quantitative Methods, Basic Assumptions; Statistics: Overview

Further Readings

Blasius, J., & Greenacre, M. (1998). *Visualization of categorical data.* San Diego, CA: Academic Press.

Greenacre, M. (2007). *Correspondence analysis in practice* (2nd ed.). London: Chapman & Hall/CRC.

Greenacre, M., & Blasius, J. (2006). *Multiple correspondence analysis and related methods.* London: Chapman & Hall/CRC.

Michailidis, G., & de Leeuw, J. (1998). The Gifi system of descriptive multivariate analysis. *Statistical Science, 13,* 307–336.

Nenadić, O., & Greenacre, M. (2007). Correspondence analysis in R, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software, 20*(3). Retrieved November 9, 2010, from http://www.jstatsoft.org/v20/i03/paper

CORRUPTION

The term *corruption* is used with reference not only to politics and to public administration but also to personal life and to business. It may refer to the perversion of any accepted standard. Common usage also includes the corruption of language, as with the corruption—that is, unauthorized alteration—of a written text. The first section of this entry examines the main problems of defining the concept of corruption. A typology of political corruption and its patterns are discussed in the second and third sections; measures, explanations, effects, and possible cures for corruption are presented in the subsequent sections; and academic analyses versus practical actions is the final topic analyzed.

Problems of Definition

Moral corruption refers to the state of an individual or an entire society in which virtues have been lost and citizens have become depraved and ruined by the pursuit of luxury and gratification. For example, what were seen as corrupt practices within the Roman Catholic Church led to Protestantism. Also, the ideals of republican virtue and the warning lessons of the moral corruption that had led to the fall of the Roman Empire were much in the minds of the framers of the constitution of the United States.

In the business sphere, corruption is said to occur when a person betrays trust and misuses knowledge or authority for personal gain, for example, if a member of a firm's staff sells trade secrets.

Political corruption involves the perversion of accepted standards of behavior in political life. In particular, it may be defined as the misuse of public office or authority for unauthorized private gain. Nevertheless, this and other definitions need to be treated with caution since the meaning of "political corruption" is itself a complex matter. The difficulty of finding a clear and agreed-on definition poses problems both for political scientists seeking to measure and compare corruption in different countries and for those responsible for drawing up anticorruption laws.

First, not all forms of misbehavior by those in positions of public trust constitute corruption. A politician who betrays secret information to an enemy country for ideological reasons is guilty of treachery but not corruption since he or she acts from an ideological motive and not for monetary gain. If that same politician reveals the secret information for payment, the action can then be characterized as corrupt as well as treacherous. Political corruption usually involves the pursuit of some form of personal gain, particularly material advantage.

Second, the mere fact that politicians or public officials derive material benefits from their positions does not automatically mean that they are corrupt. It is accepted that they are entitled to salaries and benefits for their work. It is only when they accept some unauthorized or illegal benefit that corruption is involved. However, it is not always clear which benefits a politician is or is not permitted to enjoy. For example, a major scandal in the United Kingdom (UK) ensued in 2009 from revelations about allowances used by a large number of members of parliament. In many of these cases, the parliamentarians had been permitted by the parliamentary authorities to make expense claims, which then were publicly condemned as unseemly when they were revealed. In only a few cases were the claims considered strictly illegal.

Third, and connected with the previous point, illegality is not the only criterion whereby an activity or payment may be judged to be corrupt. Especially during times of changes in public attitudes about expected standards of behavior of politicians, the law may itself be deemed to be corrupt. This is because some may see existing laws as contrary to what they feel are proper standards of morality and may wish, therefore, to alter the law.

Not only are expected standards of behavior subject to change, but also the expectations of the political class may differ from those of ordinary electors. In countries ruled by occupying or colonial authorities, there may be differences between the standards imposed by the authorities and the traditional standards within the society. Sometimes the clash of cultures is seen in the way officeholders treat members of their extended families or tribes. According to laws and rules drawn up by colonial governments or those enacted under pressure from international development aid agencies, politicians and officials must not show favoritism. By contrast, traditional social mores may oblige them to reserve jobs in the public service for their relatives. (Edward C. Banfield has referred to this traditional view as "amoral familism." Apparently, this has been an aspect of politics in Sicily. In Albania, traditional familial obligations relating to honor killings and other matters are set out in the *Kanun* [canon] of Lek Dukagjin, a code dating from the 15th century. By using the term *amoral*, Banfield is reflecting a Western view of a local culture.)

Fourth, the term *corruption* often refers not only to misconduct by those already holding elective office but also to abuses relating to the electoral process whereby such office is won.

Fifth, though corruption is used primarily to mean the unsanctioned use of public office for personal gain, it may apply too when the intended gain is not strictly "personal" but for an official's political party. For example, in a case in Philadelphia where lawyers paid bribes to obtain appointments as judges, the money went into the coffers of the ruling party and was not used for the direct financial benefit of the bribe takers.

If it is accepted that corruption is involved when the gain goes to a group (such as a political party) as well as to a politician or to the close family members of a politician, a further definitional problem arises. For it is a part of normal politics that politicians represent groups and not society as a whole. In a system where voters in a particular geographical area elect a legislator, it is expected that he or she then will try to represent their local interests. Gaining a material benefit for a locality is a normal aspect of pork-barrel politics; gaining a material benefit for a political party is a corrupt practice.

Sixth, corruption need not involve any favor by the bribe taker beyond a disclosure of unpublished information. Where the government is about to make an announcement that will affect the value of the shares of a company or sector of the stock market, prior notice enables favored persons to buy and sell shares to their advantage. Also, where a local government authority is about to make a decision about the route of a new road or rail line, a tip-off to party supporters will permit them to make property purchases or sales guaranteed to be to their advantage once the decision has been announced.

Seventh, in cases of bribery, does the corruption consist in the taking of a bribe or in the giving of an unauthorized favor in return for the bribe? If a judge accepts a bribe from a defendant but then finds the defendant guilty as charged, and the bribe did not affect the judge's decision, is the judge guilty of corruption?

Eighth, one defining test for whether an act is or is not corrupt is transparency. Typically, unacceptable transactions must be conducted in secret. A judge who accepts a bribe will be foolish to boast about it. However, not all corrupt acts are secret. Some politicians are so powerful that they feel able to flaunt their privileges. The Romanian dictator Nicolae Ceauceşcu (1918–1989) felt so confident in his immunity from prosecution that he built huge palaces for himself.

Ninth, there are differences of usage between scholars as to whether "political" corruption includes the acts of civil servants and other public employees or whether it should be restricted to elected officials and those who owe their jobs to political influence.

Types of Political Corruption

It is convenient, in the first place, to categorize corruption by the position held by the corrupted person, as, for example, with the following:

Legislative corruption

Judicial corruption

Police corruption

Administrative corruption

Corruption of local and national political executives (presidents, prime ministers, ministers, and mayors)

When there is corruption in obtaining office, it can be classified as

Electoral corruption

A second categorization is by the type of action involved. According to the U4 Anti-Corruption Resource Centre group of development agencies, it is convenient to distinguish between corrupt actions by those already in power aimed at gaining unauthorized benefits ("corrupt accumulation and extraction") and corrupt actions aimed at winning and retaining political office ("corrupt means of power preservation"). Corrupt accumulation and extraction includes the following:

Bribes, "commissions," and fees taken from private sector businesses

Undue extraction through taxation and customs

Fraud and economic crime

Politically created rent-seeking opportunities

Politically created market favors benefiting businesses owned by political elites

Off-budget transfers and manipulated processes of privatization

Extorting party and campaign funding from the state, private sector, and voters

Corrupt means of power preservation include actions such as the following:

Buying political support and majorities from other parties and politicians

Co-optation and maintenance of patron–client networks

Buying decisions from parliament, judiciary, and control and oversight bodies

Favoritism and patronage in allocation of government resources

Buying voters and votes and electoral fraud

Use of public money for political campaigns

Buying off the media and civil society

A third categorization distinguishes between national (high-level) grand corruption and local (low-level or petty) corruption. Examples of petty corruption are small payments demanded by officials to avoid delays in processing requests for licenses, passports, and other routine documents ("speed money") or small bribes demanded by

Table 1 Money Allegedly Embezzled by Leading Politicians

Head of Government	*Estimates of Funds Allegedly Embezzled (in U.S. dollars)*
Mohamed Suharto, President of Indonesia, 1967–1998	15 billion to 35 billion
Ferdinand Marcos, President of Philippines, 1972–1986	5 billion to 10 billion
Mobutu Sese Seko, President of Zaire, 1965–1997	5 billion
Sani Abacha, President of Nigeria, 1993–1998	2 billion to 5 billion
Slobodan Milosevic, President of Serbia/Yugoslavia, 1989–2000	1 billion
Jean-Claude Duvalier, President of Haiti, 1971–1986	300 to 800 million
Alberto Fujimori, President of Peru, 1990–2000	600 million

Source: Global Corruption Report 2004, Transparency International. Retrieved November 9, 2010, from http://www.u4.no/themes/political-corruption/introduction.cfm

Note: Estimates of funds allegedly embezzled are drawn from publicly available sources. The estimates do not reflect any legal developments that have occurred since 2003.

traffic police or customs officials. Grand corruption or "kleptocracy" is the demand for huge payments (typically, from those seeking major government contracts) for top politicians.

In its *Global Corruption Report 2004*, the anti-corruption organization Transparency International gave estimates of the money allegedly embezzled by some leading politicians (see Table 1).

Patterns of Corruption

Corruption not only takes many forms and exists in different parts of a political system, but it is usually clandestine and is thus hard to measure. Thus, any attempt to assess which countries are most affected and which parts of their political systems are especially corruption-prone is problematic.

The following activities frequently feature in corruption scandals in many countries.

Arms Sales

Contracts for the purchase of guns, military aircraft, and other forms of defense equipment may involve sums of money amounting to billions of U.S. dollars. Because the competition for such contracts is so fierce and the stakes are so high for the providers of the equipment, massive bribes may be given to secure them. Some of the most spectacular corruption cases of recent decades have involved the arms trade. The Lockheed Scandal of the 1970s involved bribes by the United States aircraft company to the then Japanese prime minister Kakuei Tanaka (who was sentenced to 4 years' imprisonment but remained free on appeal at the time of his death in 1993); Prince Bernhard of the Netherlands; and to leading political figures in Italy, West Germany, and elsewhere. The Bofors scandal of the 1980s concerned bribes reportedly paid by the Swedish arms manufacturer to the Indian premier Rajiv Gandhi to win a contract for howitzer guns. The Al-Yamama deals involved contracts for military aircraft between the British manufacturer British Aerospace (BAE Systems) and Saudi Arabia. These occurred between 1985 and 2006. Following press and television reports of corrupt payments to members of the Saudi royal family, there was an inquiry by the U.K. Serious Fraud Office. However, this was dropped in December 2006 under pressure from the Saudi authorities. The U.K. attorney general announced that the investigation was being halted on national interest grounds. The Agusta-Dassault Scandal led in 1995 to the resignation of Belgium's former foreign minister as secretary-general of the North Atlantic Treaty Organization and to his later conviction. Belgium's former deputy premier was assassinated. The scandal involved the purchase of military helicopters by the Belgian army.

Illegal Sales of Alcohol, Illegal Gambling, Pornography, Prostitution, and Drugs

Attempts to ban forms of vice in which large numbers of people—including "decent" citizens—indulge have led regularly to the corruption of law enforcement authorities by gangs of organized criminals. To service an extensive clientele, vice activities need to be open or semiopen. If drug takers can find the illegal suppliers, so can members of the police. To safeguard their illegal activities, the organized gangs who control them have a strong incentive to bribe members of police vice squads. In addition, they need to ensure that politicians tolerate such bribery.

Characteristically, organized criminals, therefore, supply campaign contributions to potentially cooperative politicians.

Local Government Corruption

Decisions on zoning (whether land may be used for building purposes) greatly affect the value of land. Thus, bribery of local politicians who have the authority to make or to influence zoning decisions is common. It has been the subject of major ongoing scandals in Spain. (A similar scandal led in the 1970s to the resignation of the British deputy premier.) Another common form of local bribery concerns kickbacks on local government contracts. The *tangentopolis* ("bribe city") scandal in Milan, Italy, which broke in the early 1990s, is a major example. A further common form of local government corruption is the assignment of jobs, housing benefits, and other forms of patronage as spoils to party activists. Such patronage politics was a defining feature of the "machine politics" in cities in the United States in the 19th century and in much of the 20th century.

Corruption Relating to Campaign Contributions

Political donations may be considered corrupt if

1. they contravene existing political finance laws,
2. they derive from the proceeds of a corrupt transaction,
3. they involve the unauthorized use of state resources for partisan purposes,
4. they are accepted in exchange for a promise of some unauthorized favor or contract,
5. they are from disreputable sources (such as drug barons or other criminals), or
6. the money is to be spent on illegal electoral practices such as vote buying.

There have been major scandals in recent years in Antigua and Barbuda, the Bahamas, Brazil, the Czech Republic, Croatia, Ecuador, Germany, India, Italy, Japan, Mexico, Pakistan, Papua New Guinea, Spain, Suriname, South Korea, the UK, and Venezuela.

Electoral Corruption

The stuffing of ballot boxes with fake votes, deliberate miscounting of votes (sometimes by officials who have received bribes), and vote buying are some of the most common forms of electoral corruption. Other corruption techniques include the inclusion of fake names on the voters' register and personation (falsely pretending to be another voter). Such practices are potentially widespread in elections that are held in countries wracked by violence or in developing nations where it is almost impossible to draw up a reasonably accurate register of eligible electors. In advanced democracies too—the UK is an example—the phenomenon of electoral fraud is not unknown, especially in some urban areas with mobile populations.

Measures of Corruption

It is extremely hard to measure the overall extent of corruption within a locality or an entire country. A count of the number of legal cases may do nothing more than demonstrate the effectiveness (or lack of it) of law enforcement authorities or of the press in bringing attention to scandals. Moreover, cases may vary in importance, thus making a simple count of prosecutions misleading. Any attempt to compare levels of corruption in different countries runs up against the definitional problems already described: Behavior that is considered corrupt (or is actually illegal) may not be seen in the same way in other countries.

Especially since the 1990s, various anticorruption organizations therefore have used surveys or the subjective estimates of experts as surrogate measures of corruption or of indicators of corruption (such as the degree of transparency within a country). There is considerable competition between rival corruption indices. Transparency International promotes its Corruption Perceptions Index as well as a Bribe Payers' Index and Global Corruption Barometer. The Opacity Index comes from the accounting firm PricewaterhouseCoopers. The Center for Public Integrity entered the field with a Public Integrity Index—one of the best indices in "State Capture," developed by World Bank. A total of 20 measures of transparency are listed by Ann Bellver and Daniel Kaufmann of the World Bank in their August 2005 paper "Transparenting Transparency: Initial Empirics and Policy Applications." In 2009, the Corruption Perceptions Index rated New Zealand as the least corrupt and Somalia as the most corrupt of the 180 countries studied.

One of the curious by-products of the rivalry between different anticorruption organizations is that Transparency International treats as strictly confidential the base data on which it bases its Corruption Perceptions Index. The value of these well-publicized metrics is open to question, especially since the results of information from country experts depend inevitably on the choice of experts. Moreover, there is a tendency to rely on the assessments of experts from economically advanced countries about the state of corruption in poorer countries.

Causes of Corruption

According to one view, corruption exists to evade governmental regulation. It follows that the less the regulation, the smaller the scope for corruption. Thus, socialist countries—such as the former Soviet Union—experienced an exceptionally high level of corruption as members of the power elite sought ways to evade egalitarian policies. Similarly, the prohibition of alcohol in the United States proved impractical. It merely produced widespread corruption and encouraged organized crime. This—some argue—is the inevitable consequence of laws to regulate widespread forms of private behavior.

There is an opposite view that capitalism, not socialism, encourages corruption. It is the essence of the capitalist ethic—according to this argument—that "greed is good." Hence, companies will do everything they can to bend the rules if it is in the interest of making a larger profit. Large corporations are especially rapacious when it comes to dealings in foreign, less developed nations.

A third view, popular in the 1960s and 1970s, is that corruption is especially prevalent during periods of social and economic transition. Developmental theorists such as Arnold J. Heidenheimer posited that there is relatively little corruption in socially stable, traditional societies. When there is rapid industrialization and migration, the old norms are abandoned, and corruption flourishes. However, corruption then greatly declines (without the need for any anticorruption policies) with the coming of an affluent society in which middle-class respectability and suburban lifestyles predominate.

This theory now seems much less convincing than it was at the time Heidenheimer published his influential work in 1970. The Watergate scandal in the United States during 1972 to 1974 and the Poulson Affair in the UK at about the same time showed that the theory of middle-class civic virtue was overoptimistic.

Several other causes are sometimes suggested: lack of democracy, bad governance, low salaries for public officials, one-party rule in a locality, or lack of a free and active press. Such explanations either are too vague or they are (at least partly) defective and subject to notable exceptions. For example, Singapore is a country with low democratic credentials but with a well-regarded record in the field of anticorruption.

Effects of Corruption

In the 1960s and 1970s, influential writers such as Samuel Huntington, Nathaniel Leff, and Colin Leys argued against the "moralistic" view that corruption is necessarily damaging. In fact, corruption could oil the wheels of progress by enabling investors to find ways around restrictive rules. Especially in developing countries, corruption therefore was sometimes to be welcomed. This was the predominant view of Heidenheimer mentioned previously.

There was a sudden and massive change of opinion following a World Bank report in 1969 on corruption in Africa. From then onward, corruption came to be seen as a major evil and as an impediment to economic growth. The World Bank itself became active in promoting a campaign against corruption. Transparency International was founded in 1993 as a campaigning organization, partly by former officials of the World Bank.

The sweeping statements that have been made first about the virtues and then about the vices of corruption have tended to act as substitutes for open-minded studies of the effects of corruption. The problem of relating corruption and economic development is seen by comparing countries with high levels of corruption—such as South Korea and the Philippines. In the Philippines, economic growth has been low, but it has been very high in South Korea. This shows that the effects of corruption on economic growth are, at the very least, complex and serves as a warning against generalizations and moral exhortations.

Cures for Corruption

Many different cures for corruption have been suggested, though none of them has proved to be widely effective.

The proposals include

1. cutting the number of regulations and freeing markets;
2. international anticorruption conventions, such as that of the United Nations, which came into force in 2005;
3. national laws to outlaw international bribery, such as the U.S. Foreign Corrupt Practices Act of 1977;
4. specialized anticorruption agencies such as Hong Kong's Independent Commission Against Corruption;

5. campaigns by nongovernmental organizations such as Transparency International;
6. improved accounting procedures;
7. improved procedures to ensure competitive bidding for public contracts;
8. improved international efforts to combat organized crime;
9. higher salaries for officials to reduce incentives for corruption;
10. better training for law enforcement officials;
11. ethical training for officials;
12. reforms of libel and other press laws to enable journalists to expose and report corruption;
13. aid conditionality—threatened or actual withdrawal of assistance to countries whose governments fail to implement anticorruption measures;
14. measures to combat international money laundering; and
15. reforms of laws relating to the funding of political parties and election campaigns to reduce the reliance of candidates and parties on large (and possibly corrupt) political donations.

Conclusion: Academic Studies Versus Practical Actions

In the past 20 years, international organizations, governmental agencies, and pressure groups have devoted major attention and large sums of money to the fight against corruption. This campaign has had a vital impact on the academic study of corruption. It has tended to divert attention from fundamental research into areas related to advocacy and to the recommendation of cures. Most of the corruption indices, which have become the vogue, provide pseudomeasurements. They may be of some practical value, but they are dangerous if treated with too much reverence.

Michael Pinto-Duschinsky
Brunel University
Uxbridge, United Kingdom

See also Administration; Clientelism; Election Research; Electoral Campaigns; Elites; Judiciary; Party Finance; Police

Further Readings

Banfield, E. C. (1958). *The moral basis of a backward society*. Glencoe, IL: Free Press.

Heidenheimer, A. J. (Ed.). (1970). *Political corruption: Readings in comparative analysis.* New Brunswick, NJ: Transaction.

Heidenheimer, A. J., & Johnston, M. (Eds.). (2002). *Political corruption: A handbook*. New Brunswick, NJ: Transaction.

Kaufmann, D., & Bellver, A. (2005). *Transparenting transparency: Initial empirics and policy applications.* World Bank Policy Research Working Paper. Retrieved March 30, 2011, from http://siteresources.worldbank.org/INTEAPREGTOPRURDEV/Resources/573691-1175901454225/seminar1_background_reading.pdf

Kotkin, S., & Sajo, A. (Eds.). (2002). *Political corruption in transition: A sceptic's handbook*. Budapest, Hungary: Central European University Press.

Pinto-Duschinsky, M. (2002). Financing politics: A global view. *Journal of Democracy, 13*(4), 69–86.

Robinson, M. (Ed.). (1998). *Corruption and development*. Abingdon, UK: Frank Cass.

U4 Anti-Corruption Resource Centre: http://www.U4.no

Corruption, Administrative

Administrative corruption is the abuse of roles, powers, or resources found within public bureaucracies. It may be initiated by line or staff officials, their superiors, or the agency clients. The latter will usually be private parties (e.g., applicants for licenses), but particularly in large and centralized governments; clients might also be individuals or institutions from elsewhere in the public sector. This entry considers the complexities of defining administrative corruption, some common varieties, major causal factors and consequences, the central role of administrative corruption and public administration theories in reform movements, and the effects of changing relationships between the public and private sectors on the basic concepts of administrative corruption.

Administrative corruption is a subset of the broader phenomenon of corruption, and is commonly distinguished from political corruption. Corrupt practices in the administrative realm, such as bribery, extortion, graft, patronage, and official theft, to name a few, can occur in the political arena too and, indeed, can be common in private sector administration. Indeed, some varieties of administrative corruption involve "middlemen" who station themselves at the boundaries between state and society, promising to reduce transaction costs for citizens and bureaucrats alike—even though such gains may be small or even illusory. Still, the term *administrative corruption* is usually reserved for abuses in the realm of government agencies and programs.

Conceptually, administrative corruption shares some, but not all, of the definitional problems that plague the analysis of corruption generally. Formal rules and roles are also usually more clearly specified in administration than in political life. But in societies with dominant and pervasive government agencies, or with very weak public institutions, the limits of the administration or its internal organization may not be particularly clear. Where the state lacks legitimacy or credibility—or, where it is dominated by a dictator, ruling inner circle, the military, or one political party—the formal and de facto norms of administration may diverge sharply. Moreover, what constitutes "abuse" may well be more than just a question of administrative process: Laws and regulations can be vague or contradictory (perhaps deliberately so). Obligations and preferences flowing from politics or social ties may not only cause rules to be broken but can also become normative systems themselves, rivaling formal rules in salience and power.

Like corruption generally, the causes of administrative corruption may be personalistic, institutional, or systemic. In the first category are factors such as venal, or poorly recruited and poorly trained personnel. The second includes not only poor institutional management but also administrative systems that encourage or conceal corruption because of their internal structures (consider a tax collection agency that gives individual agents the power to alter assessments as well as to collect funds) or incentive systems (e.g., very low pay). Systemic causes of corruption might include excessive political interference by elected officials or citizens, widespread poverty, or a lack of government legitimacy.

The consequences of administrative corruption have been a matter of much debate. Many have portrayed it as functional—as "grease for the wheels," providing incentives to speed up official processes—or as creating informal price systems and a rough-and-ready kind of accountability that would not otherwise exist. Since the late 1980s, however, improved theory and data have shown that administrative corruption is much more like sand in the gears. There is not, after all, a finite amount of inefficiency in administrative agencies; bureaucrats—particularly the very poorly paid officials of developing states—can create much more of it if they stand to gain as a result. Bribes are thus more powerful signals that money can be made by contriving new requirements, "losing" records, or simply doing nothing until clients pay up. Regulatory and extractive functions such as inspections and tax collections are subject to a similar logic.

Measuring corruption of any sort is problematic because it is usually clandestine and lacks an immediate victim with a stake in reporting abuses. Vulnerability to administrative corruption may be somewhat easier to estimate; however, compiling indicators of government performance and comparing them against realistic benchmarks is a promising approach. A licensing process that takes 7 weeks and involves 33 steps in City A, for example, likely has worse corruption than a similar process requiring 4 days and 5 steps in City B. The numerous steps and long delays in the first instance may well reflect past corruption—bureaucrats have learned that they can make money by adding requirements and delays, as noted above—and also tell us something about current incentives sustaining it, since the longer and less responsive the process, the greater the temptation becomes to lay out some "speed money."

Administrative corruption has generally been the main focus for anticorruption movements around the world. That is so for several reasons. At a practical level, many of the international and intergovernmental bodies that have energized reform efforts are barred from involvement in societies' internal politics. Nongovernmental organizations and grassroots reformers often find it prudent to de-emphasize political corruption, since taking on such issues often means confronting powerful entrenched interests. Focusing on administrative corruption is a way to get at the problem without directly confronting regimes and may allow reformers to draw on legal and administrative sources of support, for which there may be few counterparts in political life.

Initiatives against administrative, as opposed to political, corruption also draw on more extensive bodies of theory and evidence. In the former category are a variety of "micro"-theoretical approaches that strip processes and structures down to their essentials. Examples include principal–agent or principal–agent–client models that simplify administrative pathologies in useful ways, identifying vulnerabilities and incentive problems. Another elegant framework, Klitgaard's widely applied "equation" suggesting that Corruption = Monopoly + Discretion − Accountability, considers the individual official's power in somewhat greater detail and compares those powers against key constraints. Theories of, and arrangements for, queuing—that is, how cases are assigned to officials—highlight the importance of client vulnerabilities and official collusion, both on a single structural level and involving dealings with superiors. Often, such micro accounts of administrative corruption can be usefully nested within the broader analysis of *rent-seeking* opportunities and practices. A shortcoming of some micro theories is that they treat administrative corruption and bribery as nearly synonymous, underestimating both the variety of corrupt practices and the ingenuity of both officials and clients seeking illicit benefits.

At the overall organizational level, *public administration* theories examine structure and process, internal control, management of personnel and resources, and accountability of several sorts. Historically, the struggle against administrative corruption has helped shape and establish the value of public administration approaches. Frank Anechiarico and James Jacobs point out that in the United States, an anticorruption movement emerging out of abolitionism and determined to reform corrupt cities intertwined with emerging theories of public administration to define four distinctive "visions" of administrative corruption and reform, each dominant within a particular era. The "Antipatronage Vision" (roughly between 1870 and 1900) was aimed not only at political reform but also—critically for the theories that followed—at driving machine patronage and the influence of political bosses out of public administration. The "Progressive Vision" (1900–1933)—like its predecessor, part of a larger political movement—sought a nearly total segregation of administration from politics and did so on the national stage; it also emphasized efficiency in internal processes as both the means and the metric for reduced corruption and better government. The "Scientific Administration Vision" (1933–1970) shifted the emphasis from political reform to rigorous internal management of government guided by organization theory, pointing to a science of administration independent of the substantive competence of specific agencies. Finally, the "Panoptic Vision" (1970 onward) took advantage of new technology to institute internal controls that were so pervasive as to become problems in their own right.

Insistence on the separation of politics from administration has thus shaped both reform and public administration thinking. But it also raises problems. Few would argue for the pervasive politicization of administration, but critics suggest that reforms, with their focus on efficiency, specialization, and independent administration, may create governments that are accountable to no one—neither voters nor elected officials.

That separation of politics and administration may come under increasing strain as privatization, the rise of parastatal bodies, and an increasing (if not always appropriate) expectation that public administration be carried out in more business-like ways alter state agencies and their relationships to society. Privatization and the devolution of government functions onto corporate-style entities are often proposed as a way of eradicating administrative corruption by taking bureaucrats out of the loop: Public functionaries with nothing to sell cannot be corrupt. The result, however, may be even less accountability and responsiveness in the ways in which people are governed, as more functions move into the private sector or are conducted, even partially, by its rules. Public–private partnerships, for their part, may invite private influence—which may well be nonpartisan but will often have strong political implications—into the halls of government. Both strategies, in differing ways, mingle state with society and public with private enterprise, and they do so in a setting of weakened political accountability. Such trends not only raise practical problems, but they also suggest that the basic distinction between administrative corruption and other varieties may be more an extension of the normative doctrine that such separation *should* be maintained than of any sharp empirical boundaries surrounding the administration process.

Michael Johnston
Colgate University
Hamilton, New York, United States

See also Accountability; Administration; Administration Theory; Principal–Agent Theory; Privatization

Further Readings

Anechiarico, F., & Jacobs, J. B. (1996). *The pursuit of absolute integrity*. Chicago: University of Chicago Press.

Johnston, M. (2010). Assessing vulnerabilities to corruption: Indicators and benchmarks of government performance. *Public Integrity, 12*(2), 125–142.

Klitgaard, R. (1988). *Controlling corruption*. Berkeley: University of California Press.

Rose-Ackerman, S. (1999). *Corruption and government: Causes, consequences, and reform*. Cambridge, UK: Cambridge University Press.

Shang-Jin, W. (1999). *Corruption in economic development: Beneficial grease, minor annoyance, or major obstacle?* (Policy Research Working Paper No. 2048). Washington, DC: World Bank, Development Research Group, Public Economics.

COST–BENEFIT ANALYSIS

Cost–benefit analysis (CBA) is an accounting framework that compares benefits and costs associated with a project for the purposes of information and discussion. As discussed below, CBA is widely used by governments, and if used properly, can reduce unnecessary or inefficient government spending. Concerns in CBA considered below include whose benefits and costs should count, on what scale the analysis should be done, how to account for the value of future costs and benefits, and how to account for risk, uncertainty, and nonmarket values.

The Framework

CBA or *benefit–cost analysis* (BCA; these terms are used interchangeably) is an accounting framework for government projects. In this framework, benefits and costs associated with a public investment or decision are laid out for purposes of information and discussion. A related framework, widely used in health care but also in other areas, is *cost-effectiveness analysis*, sometimes known as *least cost planning*. In least cost planning, the goal is given, for example, to achieve a given level of carbon monoxide in a city, and the least expensive way of achieving this goal is sought.

Externalities

The rationale for government investment is externalities. Externalities are effects that are not accounted for in market operations. They arise solely from the expenses of operating markets. Air pollution, for example, is considered an externality as the polluting firm does not take into account the damaging effect of its pollution on the surrounding population. In a market system, the population could pay the firm to reduce pollution. Or the firm could pay the individuals to tolerate it, whichever course was the cheapest. Yet any such market arrangement might be too expensive to implement due to transaction costs. Sometimes the government can institute a quasi market as in cap-and-trade permits. In this case, the government fixes the quantity of pollution allowed and then sells (or gives away) rights to this quantity. These rights can then be traded. This sort of arrangement is increasingly used in fisheries to prevent overfishing. The total quota of fish allowed to be caught is set and rights are then distributed among fishermen for portions of the total quota. The decision about whether to use a quota system, or to use, say, pollution taxes or emissions limits on the basis of the available technology is the sort of decision that CBA can be used to address. More commonly, CBA is used to address questions in areas such as transportation, dam building, and, more recently, expenditures on social programs such as Head Start.

Government Use

The U.S. Federal Government widely uses CBA under presidential executive orders. Under these orders, the Office of Management and the Budget (OMB) requires the use of CBA by executive agencies and many congressional agencies as well. OMB has created a set of guidelines for the use of CBA.

CBA is also used, but less widely, by state and municipal governments. If used properly, the CBA framework can reduce poorly conceived government spending. In fact, it has already played this role to a significant extent but not to the extent possible.

What Are Benefits and Costs?

Benefits

Benefits are measured by the willingness to pay (WTP) for gains. Thus, in evaluating the building of a new road, the government might calculate the time savings to users. Then, they would estimate the value of users' time and determine a total value of time saved. This figure would then represent an estimate of the users' WTP and therefore of their benefits.

Costs

Costs are the amount that those bearing them would be willing to accept (WTA) in order to bear the burden of the project or decision. Thus, in building a new road, the analyst would seek to determine what payment the taxpayers would have to make. In addition, homeowners who live near the road might bear certain costs in terms of pollution and noise. The amount it would take to compensate these homeowners would be an addition to the costs. Alternatively, the government might build a barrier to reduce noise so that the cost to nearby homeowners is small and might be ignored.

Standing

Not everyone affected nor all sentiments created are included in practice in a CBA. The decision about whom to include or exclude in a CBA is to decide the *standing* of those affected. Most analyses are done from a particular viewpoint. A municipality might consider only the effects to municipal revenues or only those felt within the city. The state government might not consider the effects of its pollution policy on neighboring states, and the federal government might not consider its effects on foreign countries. Many projects will have only minor effects on certain groups, and it may be considered too expensive to calculate these. In addition, certain sentiments may be disregarded as not being acceptable; for example, one might suffer harm from envy, but this harm would generally go unrecognized in a CBA. A more useful example concerns the value of stolen goods to the thief. Clearly, they have value to the thief, but in considering whether or not these goods should be returned, one would not count their value to the thief, as holding such goods is illegal.

Partial and General Equilibrium Analysis

How many markets will a project affect? Suppose a buyer is considering whether or not to buy a less expensive regular refrigerator or a more expensive energy-saving refrigerator. The energy-saving model will save energy costs each year, and the buyer wonders if the extra expense is worthwhile. This can be determined by the buyer's own personal CBA. Any effect on markets is negligible and can be ignored. An analysis such as this that is very limited in its inclusion of markets is called *partial equilibrium analysis*. However, if one is considering a countrywide policy to subsidize energy-saving models, the analysis should consider markets more broadly. When the effects in many markets are taken into account, the analysis is *general equilibrium analysis*. Consider a national oil tax and all of the markets affected: transportation markets; the price of products made with oil, such as plastics; labor markets; and others. One analysis is that such a policy should be implemented from a general-equilibrium perspective.

Discounting

Consider the choice between a more expensive energy saver refrigerator and a less expensive one. Suppose the buyer faces the following costs and benefits over time. For simplicity, we will assume that the life of each refrigerator is 5 years and that the energy costs are constant (see Table 1).

The buyer would spend an extra $200 now to save $255 over 5 years. Is it worthwhile buying the energy saver? The savings of $45 per year amounts to a total of $225, which is greater than the $200 extra cost, but this comparison fails to take into account the time value of money. Suppose you know that you can obtain a 5% yearly return from investing in a safe Certificate of Deposit (CD). The question now is what sum of money you could invest today at 5% that would yield $45 per year

Table I Discounting: An Illustration

Choice	*Initial Costs ($)*	*Energy Costs ($)* *Year 1*	*Year 2*	*Year 3*	*Year 4*	*Year 5*
Cheaper refrigerator	1,000	60	60	60	60	60
Energy-saver refrigerator	1,200	15	15	15	15	15
Cost difference	−200	45	45	45	45	45

for 5 years. The answer to this question is the *present value* (PV) of the stream of 5-year payments. This PV will be given by the formula for an annuity (a uniform stream of payments), which is

$$PV = A[1 - (1 + r)^{-t}]/(1 + r)^t,$$

where A is the yearly amount, t is time, and r is the discount or interest rate (here 5%). In the case of the energy saver refrigerator, this will be

$$PV = \$45[1 - (1.05)^5]/(.05) = \$194.83.$$

Since this PV is less than the $200 cost, it is not worthwhile buying the energy saver. If one's discount rate were, for example, 1%, then the PV of the energy saving would be $218.40, and it would be worthwhile. The lower the discount rate, the greater is the PV of future benefits. The net present value (NPV) is the PV of benefits minus costs. The CBA test is to determine if the NPV is positive, in which case it is said that the project is financially desirable. The NPV in the present example would be as follows:

$$NPV = \$194.83 - \$200 = -\$5.17.$$

Since this is a negative value, the project is not financially desirable. One way to look at this example is that if one receives less than the 5% return, one can gain from investing the $200 in a safe CD. The NPV is not the only measure used but is the one most used by economists.

Risk and Uncertainty

CBA almost always involves considering cash flows in the future. The future is always uncertain and risky. *Risk* is usually defined as occurring when the probabilities of events are known. *Uncertainty* occurs when the probabilities are unknown or subject to unknown variances. In considering the purchase of the refrigerator, for example, we might expect the energy savings to increase over time with some certain probability. This would be a question of risk. As a simple procedure, we might calculate the expected value of the future energy savings, given the known probabilities of increases in energy prices. The expected value is the energy savings times its probability. Suppose we expect the value of the energy savings to be $20 per year due to the increase in energy costs and that this increase in savings is expected with 90% probability. The expected value of this in each year would be $5.00 times 90% or $4.5 dollars per year. Taking the PV of this annuity at 5% for 5 years, we have $19.48. When we add this to the PV in the previous scenarios with the 5% discount rate, we would have a total of NPV = $14.31. Since this is positive, the energy saver passes the CBA test, given that the energy costs go up by $5 per year with 90% probability.

In a more sophisticated analysis, we would have a distribution of probabilities. The NPV can also be calculated in these situations. A common procedure is to use a Monte Carlo simulation program that allows us to account for the entire probability distribution. Further discussion of this issue lies beyond this essay.

Nonmarket Values

Although CBA is a sort of financial accounting, it considers many values not always counted in business accounting. Many preferences can be valued, although they are not always readily apparent. The value of a house with a view compared with the same house without a view is readily determined, for example. The value of cleaner air and water can also be determined. The value of fishing in a less polluted stream as compared with a more polluted stream is more challenging. Similarly, the value of visiting a national park requires more sophisticated work. Economists have developed a number of techniques for determining environmental and other values. For example, they ask how far a person might drive to arrive at a park destination. Valuations can be made by looking at the cost of driving to drivers arriving from various estimated distances from the park. Also, questionnaires can be used to determine values. Analogies from market behavior can be applied to nonmarket behavior. Prices used to value nonmarket behavior are called *shadow prices*. Thus, CBA attempts to perform a sort of social accounting to determine the best public investments.

Richard O. Zerbe Jr.
University of Washington
Seattle, Washington, United States

See also Economic Policy

Further Readings

Boardman, A. E., Greenberg, D. H., Vining, A., & Weimer, D. (2001). *Cost-benefit analysis: Concepts and practice*. Upper Saddle River, NJ: Prentice Hall.
Gramlich, E. M. (1990). *A guide to benefit-cost analysis*. Englewood Cliffs, NJ: Prentice Hall.
Lesser, J. A., Dodds, D. E., & Zerbe, R. O. (1997). *Environmental economics and policy*. Reading, MA: Addison-Wesley.
Sugden, R., & Williams, A. H. (1978). *The principles of practical cost-benefit analysis*. New York: Oxford University Press.
Zerbe, R. O., Jr., & Dively, D. (1994). *Benefit-cost analysis in theory and practice*. New York: HarperCollins College.

Credible Commitment

Individuals are generally assumed to choose rationally in ways that make themselves better off. However, there are times when a political actor can benefit by *denying* himself or herself, early in a political interaction, the option of choosing self-serving actions later on. A general may know that retreat could be a valuable option in the future, but "burning one's bridges" may elicit more advantageous behavior from one's soldiers (or one's opponents). Constraining self-interested choice to induce beneficial actions from other players is called "credible commitment."

This entry examines several systematic and foundational credible commitment problems. In international relations, credible commitment can strengthen deterrence. Further, opportunism by government officials can intimidate economic actors; healthy economic growth requires the government to be credibly constrained from opportunistic behavior. Constitutions, courts, independent central banks, and international agencies can facilitate the kind of credible commitment that enhances economic development. In democracies, credible commitment necessarily entails a limitation on the scope and power of majority rule.

Credible Commitment, Programming, and Delegation

At the heart of the problem of credible commitment is the strategic interdependence of two or more actors making sequenced choices. An example explored in the early 1960s by Thomas Schelling is nuclear deterrence. A stylized version of this problem is as follows: Assume that the Soviet Union moves first, choosing whether or not to invade Europe. The United States then responds to an invasion with either conventional or nuclear weapons. The Soviet Union's choice depends on what it thinks the United States will do in response. The Soviet Union's dominance in conventional weaponry would give it a victory in Europe if the United States restricts itself to conventional weapons. However, if the United States responds with nuclear weapons, the result will be a nuclear holocaust that is the worst outcome for both sides.

The United States would like to convince the Soviet Union of its willingness to use its nuclear arsenal. However, after an invasion, the self-interested choice for the United States is *not* to initiate a nuclear exchange. The self-interest of the United

States diminishes the deterrent effect of its nuclear arsenal. Based on its assessment of the U.S. self-interested response, the Soviet Union can invade with impunity. Paradoxically, the option of making a self-interested choice works against the U.S.'s own best interests—thus, encouraging the very action that it would like to deter.

The United States would be better off if it could *commit* itself, prior to the invasion, to a course of action that would *not* be in its best interests later on. If it could commit itself to using the nuclear deterrent, willingly initiating a nuclear war in response to a Soviet invasion, then the best choice for the Soviet Union would be *not to invade*.

But how could the United States commit itself, prior to an invasion, to an action that would defy its self-interest in the event? One answer is *programming* of behavior. If the United States creates bureaucratic machinery that mechanically implements a nuclear response in the event of an invasion, then the Soviet Union would choose not to invade.

Closely related to programming is delegation. If the response to a Soviet invasion were transparently and credibly delegated to someone who could be expected to *choose* a nuclear exchange, then the Soviet Union would be given pause. For example, the chairman of the Joint Chiefs of Staff from 1961 to 1965 was General Curtis LeMay. LeMay had not only implemented the firebombing of Tokyo during World War II, but he had also expressed support for the use of nuclear weapons. The Soviet government's awareness of LeMay's central role in strategic decision making might well have made the nuclear deterrent more credible—not because he was representative of other public officials but *because* he was unrepresentative.

The delegation solution to credible commitment problems defies our conventional wisdom. Conventional wisdom (and a good deal of principal–agent theory) concludes that selecting an agent with preferences similar to the principal's preferences is desirable; or failing this, the principal should shape the agent's behavior through incentives that cause the agent to choose as the principal would in any decision situation. However, delegation is an effective means of credible commitment only when the person receiving the delegated authority has preferences *different* from the person doing the delegating.

Political Moral Hazard and Central Banking

In the case of nuclear deterrence, credible commitment had the potential to help the United States while depriving the Soviet Union of a victory. Other uses of credible commitment make *all* players better off. For instance, consider an investor who has hidden some money underneath her mattress. If she were to invest in government bonds, then she could increase her wealth, and the government would have funds necessary to pursue its agenda.

Is the investor's money safe with government bonds? They may be a safe investment if the government pursues a stable, noninflationary monetary policy. However, democratically elected public officials inevitably have an incentive to give a surprise "boost" to the economy with monetary policies that would end up with inflation—diminishing the value of the bonds. If the government is not credibly constrained from engaging in inflationary policies, then the investor may seek a hedge against inflation and pass up the government's bonds. Both the investor and the government would be better off if monetary policy were delegated to an agency that is guaranteed to be beyond the control of the democratically elected officials—a central bank.

Creating an independent central bank is increasingly thought of as an essential feature of modern political economy—precisely because it denies public officials the opportunity to pursue their morally hazardous preferences. The evidence has shown, as noted by Robert J. Franzese Jr. (1999), that independence in central banks does decrease inflation, by differing the amounts depending on trade openness. Most interestingly, as Giandomenico Majone (1999) observes, to resolve credible commitment issues, "the delegate is chosen precisely because his or her preferences do not mirror those of the delegating authority" (p. 69).

Credible Commitment to Nonmajoritarian Institutions

The case of central banking is one in which we are accustomed to thinking of delegation to a nonmajoritarian institution as a form of credible commitment that protects the public from the dangerous impulses of elected officials. However, delegation may play a similar role in a variety of policy arenas in which elected officials experience temptations to use power opportunistically in ways that endanger property rights and contract enforcement and, thus, diminish investment and long-run economic growth.

Many of a government's actions are crucial to the strategic thinking of investors, entrepreneurs, homeowners, and other citizens. Citizens may have reason to suspect opportunistic behavior on the part of the government; this is seen around the world in the concept of "political risk." If a firm invests in an expensive, immobile manufacturing plant, will the government then use its taxation powers to extract all the profits? If an investor embarks on an investment strategy that assumes low inflation rates, will the politicians undermine that strategy with monetary surprises? If an employer hires workers assuming fixed employment regulations and contracts, will those be suddenly changed as a result of the latest election results?

All of these may be thought of as credible commitment problems, in which the prior constraint of political actors may encourage beneficial economic growth. Seen through the lens of credible commitment, a number of constitutional features can be thought of as aiding the resolution of credible commitment problems. A Madisonian separation of powers is beneficial because it diminishes the ability of the branches of government to unite in opportunistic schemes.

Separation of powers also strengthens the independence of the judiciary. In many regimes, the courts may do a good job of arbitrating disputes between citizens, but in disputes between citizens and the state, the courts may defer to the state and its authority. When economic actors are convinced that the judiciary can serve as a check on government officials, they are more willing to make investments contributing to economic development.

The creation of an independent regulatory agency may be thought of as an attempt to limit the opportunistic actions of politicians. Key decisions are delegated to commissioners and staff who have only tenuous links to political factions, and they are expected to embed their decision-making process in a fixed procedure that discourages abrupt changes in policy. In the United States, for example, the National Labor Relations Board has embedded policy decision making in a bureaucratic and legalistic process that has resulted in the long-term development of a body of labor laws that diminishes uncertainty surrounding employment contracts. In these and other cases, "Political property rights—the rights to exercise public authority in certain policy areas—have been altered by delegating important powers to nonmajoritarian institutions" (Majone, 2001, p. 58).

In general, problems of credible commitment constitute a change of direction for rational choice theories of politics. Whereas early rational choice theories sought to explain political phenomena in terms of the self-interested choices of actors, the notion of credible commitment tries to explain certain political phenomena (central banks, independent judges, and autonomous bureaucrats) as institutional manifestations of the need to *deny* rational actors opportunities to pursue their morally hazardous short-term incentives.

Problems of credible commitment constitute a challenge to democratic theory as well. While democratic theorists may think of accountability to elected officials as the sine qua non of constitutional arrangements, empowering democratic forces to make monetary and regulatory decisions can pose a threat. An alternative constitutionalist perspective maintains the necessity of dividing, checking, and ultimately constraining democratic institutions.

Gary J. Miller
Washington University in St. Louis
St. Louis, Missouri, United States

See also Bureaucracy; Central Banks; Judicial Independence; Principal–Agent Theory

Further Readings

Franzese, R. J., Jr. (1999). Partially independent central banks, politically responsive governments, and inflation. *American Journal of Political Science, 43*(3), 681–706.

Majone, G. (2001). Nonmajoritarian institutions and the limits of democratic governance. *Journal of Institutional and Theoretical Economics, 157,* 57–77.

Miller, G. J. (2000). Above politics: Credible commitment and efficiency in the design of public agencies. *Journal of Public Administration Theory and Research, 10,* 289–328.

Moe, T. (1987). Interests, institutions and positive theory: The politics of the NLRB. *Studies in American Political Development, 2,* 236–299.
North, D. C., & Weingast, B. R. (1989). Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History, 59,* 803–832.
Schelling, T. C. (1960). *The strategy of conflict.* Oxford, UK: Oxford University Press.

Crisis

At first sight, political crises appear as sudden and acute disruptions in the "normal" working of political systems. This ordinary perception of political crises is hereafter taken as their provisional definition, which offers at least the convenience of illuminating a paradoxical situation. The history of our societies is extremely rich in periods or events that may serve as examples. However, political crises have only belatedly been approached as phenomena deserving to be systematically analyzed and explained for their own sake, worthy of being elevated to the status of explicit research objects. Indeed, in most works dealing with them, political crises have been approached as a means of understanding or explaining something else. This orientation of the scholar's interest has constituted a long-lasting handicap for the knowledge of these critical phenomena. The main root of this handicap lies in the common fascination with outcomes of historical processes, especially with the outcomes of critical processes, and in the automatic assumption that these outcomes should constitute the explicanda of these process analyses. This entry examines, first, how fascination with outcomes pervades most of the approaches that have dominated this field of research; second, the perspectives that have posited crises as particular stages on the paths of political development and the ones that have attempted to insert into crisis analysis the activities and choices of actors; and third, the approaches that have aimed at identifying, within the historical paths or sequences, turning points or branchings that are supposed to explain their course toward diverse outcomes.

It is precisely the departure from this fascination with outcomes that has cleared the way for a radical reorientation of the analysis of political crises. The next part of the entry focuses on the crux of this reorientation, which consists in grasping crises as particular states of the structures of the societies affected by these crises. This perspective not only allows us to account for a considerable number of the properties of crises but also to understand why the mundane idea that analysis and explanation of periods perceived as abnormal would require exceptional approaches is misleading. The final section of the entry comprises a brief discussion of several implications of this reorientation regarding some central features of the processes of crisis, particularly the uncertainty that characterizes such processes, the frequent emergence of charismatic leaders, the question of the alleged prevalence of hidden arenas over what is at stake in the open, and, finally, the question of the autonomy of the processes of crisis from their etiology.

The Focus on Outcomes

The intellectual attraction of the outcomes of crises is perfectly exemplified by conceptualizations of political development or political modernization. For instance, in a seminal set of studies, Leonard Binder and associates have distinguished several types of crises of political development such as identity, legitimacy, participation, distribution, and penetration crises. These crises have been thought of as functional requisites or necessary conditions for the emergence of a developed political system. Concurrently with their teleological biases, such approaches, as Sydney Verba has noticed, did not succeed in distinguishing crises as processes of "disruption" of the ordinary working of political systems from problems that political elites had to solve in order to modernize their country—even though some of these problems have occasionally given birth to conflicts more closely fitting the provisional definition of crises. This difficulty is heightened in those approaches that tried to link specific outcomes (typically democracy) with their preconditions seized through global economic or social indicators (e.g., literacy or urbanization rates, gross national product per capita), because the quest for statistical correlations between these indicators and their assumed outcomes excludes from the research agenda all the main components of political crises—the activity of the protagonists as well as what shapes or constrains it.

In contrast, other approaches, while assigning themselves the same explicanda (the outcomes of crises), have attempted to rehabilitate actors' activity and strategies. Their major contribution lies in the historical description of the elites' or leaders' moves, of their tactics and dilemmas of coalition building, and of the bargainings among the "great" protagonists of the events scrutinized, particularly the bargainings of the contested regimes' incumbents with their opponents. Often rich in insights, such approaches nonetheless analyze the actors' tactical activity with the aim of identifying, for each type of outcome, a specific historical path or trajectory—especially a distinctive sequence of typical stages (each type of outcome being thus characterized by a distinctive path). This affects works dealing with the establishment of democratic systems and with the crises, collapses, and breakdowns these systems experienced as well as works concerning the processes that resulted in various types of authoritarian regimes (particularly in the interwar period in Europe) and, later, in the transitions to democracy. The intellectual appeal of *outcome* is easy to understand: Once it has occurred—but only then—the outcome (the seizure of power by Italian Fascists or the survival of democracy in France after the May 1958 crisis) matters; it often weighs heavily on the fate of individuals and societies. The appeal of outcome constitutes, however, a huge trap. First, it determines the scholar's modus operandi—as a matter of fact that of the natural history of revolutions, according to Crane Brinton: The aim being to explain the actual outcome, it seems self-evident to cut off the historical sequence through a regressive analysis, taking as a starting point the outcome itself; the selection of facts then operates so as to make them converge toward the outcome that occurred. Second, a heavy historicist bias informs such reasoning, leading the scholar to ignore de facto the *contingency* of the crises' outcomes and also implying that outcomes reflect or reveal the nature or essence of the critical processes or events that brought them about.

Many scholars have thought that they could avoid these pitfalls by introducing into historical paths some *branchings*, or moments when "history hesitates"—that is, by proposing a treelike image of the course of history. In an influential essay on the breakdowns of democratic regimes, Juan Linz has thus described the supposedly characteristic path of this particular set of outcomes by proposing roughly the following sequence: emergence of unsolvable problems/loss of incumbents' power/power vacuum/transfer of power to the antidemocratic opposition, civil war, and re-equilibration of the democratic system—here lies the crucial branching point. Remarkably enough, that way of thinking about historical branchings permeates works referring to the most different sociological traditions, such as, prior to Linz's essay, the structural Marxist analysis that Nicos Poulantzas devoted to the historical paths that led to the fascist regimes in Italy and Germany (the author combines the branchings with the idea, rather implicit in other works, that, at some stage, the processes under study reach points of no return and are conceived of as reversible before that point). Unfortunately, the remedy—the introduction of branchings—proves insufficient; it even leads to a dead end. Although these approaches claim to emphasize the role of political leaders and their decisions, they retain the hard core of historicist reasoning. They link each type of outcome—whether the process led to the cooptation of disloyal opposition, to civil war, and so on—with one, *and only one*, historical path supposed to be characteristic of it and supposed to be different from the paths or sequences ascribed to processes that resulted in other types of outcomes. We shall go back to this question again below from a different viewpoint. Nevertheless, it is far more meaningful that, commenting on their own analysis of transitions to democracy, Guillermo O'Donnell and Philipp Schmitter—who took Linz's sequence of stages, supposedly typical of the breakdown of democracy, as an explicit model of explanation—acknowledged their failure to uncover a distinctive historical path proper to transitions. But this failure itself constitutes an interesting result of their work—probably the most important one. This observation actually applies to all the approaches mentioned previously.

Crisis as Structural Transformation

There is only one way to elude what is merely a particular kind of belief in laws of history (laws of historical development): namely, to bracket off, at least temporarily, the outcomes of crises and, more broadly, the outcomes of critical processes and to break with the idea that to account for these phenomena consists in explaining their outcomes—that is, to choose a different puzzle to solve. But what is at stake in this move is not only the elimination of historicist biases; it is above all to identify and explain what the crises are made of and what happens during crises. Everyone today agrees on the idea that matter—physical matter—exists in different states: solid, gaseous, or liquid. For the theory of fluid conjunctures outlined by Michel Dobry, social structures and, more broadly, configurations of institutions and social relations, even when deeply objectified, are not necessarily more structurally stable than is matter. They can undergo short-term transformations of their state, and they can experience different states (which gain from being termed *conjunctures*, as particular states of social structures). This perspective thus opens the possibility of bringing to light the plasticity of structures and their sensitivity to actors' mobilizations.

All this must be specified. First, the theory of fluid conjunctures applies exclusively to critical conjunctures peculiar to complex social systems. These systems are defined by their high internal differentiation into a multiplicity of social sectors, spheres, or fields, relatively autonomous from each other, strongly institutionalized, tending toward self-reference, and each of them endowed with its own social logic (a set of constraints, resources, stakes, pragmatic and normative rules of the game, etc.), a logic that weighs heavily, in routine conjunctures, on the perceptions, calculations, and tactical activities of actors. This means that within these systems—which are not necessarily democratic ones—ordinary politics operates as a relatively autonomous sector or field, with its own specific logic. Second, many of the empirical regularities of the political crises that affect these systems have been already noticed or observed by their actors or by scholars. The point is that these regularities can be fully explained when thought of as a process that alters, in a short term, the structural features of these systems, which, in the first place, is their particular form of differentiation—that is, more precisely, when thought of as a process, or as a dynamics, of desectorization of these systems. Crucially, this dynamics—which occurs particularly under the impact of multisectoral mobilizations (those that spread, on a competitive mode, in several differentiated social sectors)—is a process that accounts also for a lot of regularities previously poorly identified or unidentified.

This is the main reason why, to understand political crises or, better, fluid conjunctures, the major features of this desectorization dynamics must qualify—or even substitute for—the provisional definition sketched above. Desectorization first involves a manifest reduction, and sometimes the complete collapse, of the autonomy of the different sectors; a visible fading or crumbling of sectoral frontiers; and a rapidly growing permeability to tactical moves and to stakes external to sectoral social logics, notably external to those of the political sectors or fields. Among other features, this results in a dramatic and unexpected mobility of stakes—puzzling for protagonists of crises as for observers—sudden disruptions of sectoral institutionalized temporalities or rhythms, rapid alterations of the ordinary division of political labor, and the no less abrupt outbursts of processes of de-objectification—the loss of the natural, taken-for-granted dimension that is also the legitimacy—of the social reality of institutions and collective bodies.

Further, the effects of this conjunctural desectorization can be fully checked when looking at the perceptions, anticipations, tactical dilemmas, calculations, and moves of actors. This deserves, however, a detour by one delicate question: the belief that abnormal or exceptional periods require an exceptional approach, different from the viewpoint of "normal" social sciences. Let us revert to the historical branchings discussed previously: In most works, the introduction of such branchings cumulates the effects of historicism and the effects, a priori logically incompatible, of methodological exceptionalism. Ultimately, justified by the alleged abnormal—in olden times called "pathological"—characteristic of crisis phenomena, this exceptionalism takes a classical form. On the one hand, periods of political stability, or political "equilibrium," would fall within the scope of so-called normal, that is, structural (whatever meaning one gives to this notion), and deterministic approaches. On the other hand, periods of crisis (or "transition," "revolution," etc.) would require, by their very nature, to be explained in terms of choice, of decision making, or of strategic calculation. Such an exceptionalism, seemingly taken for granted, is constitutive of the heroic fallacy. First, it leads the scholar to abandon any sustained interest in structures and for what they become during critical periods. Second, it conveys a conception of actors' choices or decisions as, at a minimum, underdetermined and, most often, undetermine—that is, free from the causalities that "normal" social sciences attempt to establish. It is noteworthy that this fallacy is not confined to scholars: It replicates the ordinary point of view of practitioners of modern revolutions (e.g., Lenin or Trotsky) who, to cope with revolutionary situations, oppose the categories of objective factors and subjective factors of revolutions and conceive of situations such as the exclusive realm of subjective factors. All these remarks entirely apply to the standard conception of critical junctures: If these junctures are supposed to constitute the starting points of path dependency processes, they are usually thought of as moments of choices, choices freed from any determination or structural constraint; that is why critical junctures—which can also, as many authors forget and as Paul Pierson reminds us, correspond to events in no way critical—must not be confused with fluid or critical conjunctures. The heroic fallacy presumably is an illusion necessary to the practice of actors. In terms of knowledge, it leads to losing at once on several counts: first, by ignoring what structures become in periods of crisis and, second, because it excludes any possibility of identifying and explaining not only the factors shaping actors' perceptions and actions but also the differences distinguishing critical conjunctures from routine ones. This has an important consequence: We have to refuse to change approaches when switching from the analysis of normal periods to the analysis of abnormal or exceptional periods. In other words, the understanding of political crises calls for a *methodological normalization* of these phenomena (i.e., for a principle of methodological continuity).

The question especially matters when confronting the uncertainty that the actors of crises have to deal with. This uncertainty does not boil down to a narrow uncertainty on the outcome of crises or on the postcrisis course of possible ensuing path dependency processes—this is only a marginal part of it. Uncertainty, rather, is an overwhelming, multiform, and pervasive component of the situational constraints stemming from the desectorization of social space, which affects all the protagonists of crises. The specific social logics of the various differentiated social sectors, spheres, or fields cease to be the reference and the material of actors' calculations, assessments of varied lines of action's efficiency, and definitions of situations. The often dramatic collapse of the sectoral logics' hold on actors' calculations goes hand in hand with the emergence of a widened form of interdependence that tends to replace social games—particularly political games—that, in routine conjunctures, are characterized by more local, sectorized, and compartmentalized forms of interdependence. This widened interdependence not only tends to introduce into actors' calculations resources, actors, and stakes that are external to the ordinary logics of political sectors but also deprives actors of their routine means or tools of calculation and anticipation and their routine references, clues, markers, pragmatic rules of the game, and so forth. Widened interdependence is the mainspring of the sudden and unpredictable variations of the efficiency, or value, of various lines of action and resources. It simultaneously results in making it difficult and often very costly for actors to access information on that efficiency as well as information on orientations of other actors and, regarding collective actors, on their very consistency and identities (Who exactly are my allies today? Who are the real hardliners? Are the moderates still moderates?). That is why it is relevant to term this uncertainty *structural uncertainty*, in its strong meaning. This entails that conjunctures of crisis are no less constraining for actors than are normal situations; actors are no less rational, but they are compelled to manage to calculate, anticipate, and orient themselves in other ways. For instance, from this derives, in these conjunctures, the special and growing attraction of situational focal points.

Implications for the Study of Crises

All this has multiple and varied implications. It first induces a reconsideration of several classical *topoi* of political sociology. Two of them can illustrate the point fairly well:

1. The Weberian theory of charismatic leadership in critical periods stresses the personal qualities, words, claims, and deeds of the leader; however, as in the case of the French May 1958 crisis, this conception should be questioned: Actually, the charismatic leadership attributed to General de Gaulle stems primarily from the fact that he represented a strictly situational focal point for diverse opposing actors, most of whom were not his proponents.

2. Delegitimization processes in political crises are traditionally conceived as necessary preconditions of crises, but the desectorization dynamic peculiar to fluid conjunctures (and the resulting de-objectification, or de-reification, of institutions) prompts us to think of delegitimization processes also, and often primarily, as by-products of desectorization dynamics, and to observe that crises may occur without any prior delegitimization of authorities or regimes.

Second, it permits accounting for usually misinterpreted empirical phenomena—for example, the idea that the hidden overdetermines and the visible informs many analyses of critical processes. In this way, while somewhat reticent about conspiracy theories, Linz nevertheless asserts that in one of the crucial stages of the breakdown of democratic processes, we should observe a narrowing of the political arena, marked by the regime leaders' opening up to the disloyal opposition (e.g., Mussolini in the Italian case), the supposedly decisive role of secret bargaining, and the replacement of the parliamentary arena by a hidden one, accessible only to a very restricted number of individuals. Although in such contexts—as in many others, often more ordinary ones—actors may be tempted by secret negotiations and moves, we hardly find this kind of overdetermination of the visible or the open by the secret. Since these critical contexts are situations of widened interdependence, hidden arenas, restricted groups, conspiracies, and covert agreements per se can never benefit from any causal or ontological superiority over visible moves and events and, more specifically, over the expressive, nonintentional information that visible moves carry and that focus all actors' attention, whether they like it or not. The March on Rome, seen by some historians as an inefficient theatrical staging, nonetheless constrained, at that moment, the anticipations, calculations, and tactics of all protagonists, even those operating undercover. Moreover, Linz rightly points out the coming into play, during that stage, of actors external to the political field, such as the church, the unions, and the army or high-ranking civil servants; obviously, all of this is much more consistent with the desectorization perspective than with the narrowing hypothesis.

There remains one last implication to consider. According to Carl von Clausewitz, if war is the continuation of political relations by other means, it nevertheless has the property, once produced by politics, of depending also on its own dynamics or its own "grammar," which is the rise to the extremes of violence, whatever the ends or will of actors. The reader already knows that the dynamics of political crises, despite the imagery of the intensification of conflict and of the rise in the scale of violence haunting many works, is not that of wars. And yet, like wars, political crises tend to escape, to pull themselves away from their conditions of genesis, but because of their own distinctive dynamics, that is, that of desectorization. The autonomy of critical processes from their causes or conditions of emergence is not innocuous: It further undermines the premises and causal imagery of historicist approaches and prompts a critical reflection on what the social sciences have often expected, and still expect, from the unveiling of the historical conditions of emergence of political crises and critical processes—particularly the belief that the ultimate sociological truth of these phenomena exclusively or primarily lies in the factors that engendered them. It is easy to guess that the target of this criticism is not causality as such but rather some pedestrian forms of misplaced determinism.

Michel Dobry
Université Paris 1 Panthéon-Sorbonne
Paris, France

See also Breakdown of Political Systems; Charisma; Conflicts; Democratization; Legitimacy; Revolution; Transition

Further Readings

Almond, G. A., Flanagan, S., & Mundt, R. (Eds.). (1973). *Crisis, choice, and change: Historical studies in political development*. Boston: Little, Brown.

Berg-Schlosser, D., & Mitchell, J. (Eds.). (2000). *Conditions of democracy in Europe, 1919–39: Systematic case studies*. London: Macmillan.
Binder, L., Coleman, J. S., LaPalombara, J., Pye, L. W., Verba, S., & Weiner, M. (1971). *Crises and sequences in political development*. Princeton, NJ: Princeton University Press.
Brinton, C. (1965). *The anatomy of revolution.* New York: Vintage. (Original work published 1938)
Dobry, M. (Ed.). (2000). *Democratic and capitalist transitions in Eastern Europe: Lessons for the social sciences.* Dordrecht, Netherlands: Kluwer.
Dobry, M. (2009). *Sociologie des crises politiques: La dynamique des mobilisations multisectorielles* [Sociology of political crises: The dynamics of multi-sectoral mobilizations]. Paris: Presses de Sciences Po.
Linz, J. J., & Stepan, A. (Eds.). (1978). *The breakdown of democratic regimes.* Baltimore: Johns Hopkins University Press.
Moore, B., Jr. (1966). *Social origins of dictatorship and democracy.* New York: Oxford University Press.
O'Donnell, G., & Schmitter, P. C. (1986). *Transitions from authoritarian rule: Tentative conclusions about uncertain democracies*. Baltimore: Johns Hopkins University Press.
Pierson, P. (2000). Path dependence, increasing returns, and the study of politics. *American Political Science Review, 94*(2), 251–267.
Poulantzas, N. (1974). *Fascism and dictatorship: The third international and the problem of fascism* (J. White, Trans.). London: New Left Books. (Original work published 1970)
Schelling, T. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Crisis Management

We speak of crisis when a community of people—an organization, a town, or a nation—perceives an urgent threat to core values or life-sustaining functions, which must be urgently dealt with under conditions of deep uncertainty. Crisis management pertains to all activities addressing that threat and aimed at minimizing its consequences.

Crisis and the Modern Society

Public authorities face a variety of crises, such as natural disasters and war, financial meltdowns and terrorist attacks, epidemics and explosions, and infrastructural dramas and failures of information and communications technology (ICT). Crises are not routine events (e.g., fires or traffic accidents). Crises are inconceivable events that often take politicians, citizens, and the media by complete surprise.

Crises create tough challenges for public authorities and their organizations. Citizens expect public authorities to safeguard them from the threat, make critical decisions, and implement them—under considerable time pressure and in the absence of essential information about causes and consequences. Two factors make it increasingly hard to meet these challenges.

First, the very qualities that increase welfare and drive progress in modern societies make them vulnerable to crises. Modern society has become increasingly complex and integrated. Complexity makes it hard to fully understand emerging vulnerabilities, which, as a result, can go long unrecognized; attempts to deal with them often produce unintended consequences (fueling rather than dampening the crisis). Tight coupling between a system's component parts and those of other systems facilitates the rapid proliferation of disturbances throughout the system. Crises may thus have their roots far away (in a geographical sense), but they can rapidly snowball through global networks, jumping from one system to another and gathering destructive potential along the way.

All this makes it very hard to recognize a crisis before its consequences materialize. When a crisis begins to unfold, policymakers often do not see anything out of the ordinary. Hidden interactions eat away at the pillars of the system, but it is only when the crisis is in full swing that policymakers can recognize it for what it is. Once a crisis has escalated into view, authorities can only try to minimize its consequences.

Second, the contested nature of crisis complicates the situation. A crisis rarely, if ever, "speaks for itself." The definition of a situation is, as social scientists say, the outcome of a subjective process and is continuously subjected to the forces of politicization. One person's crisis is another's opportunity. For public authorities, this spells trouble: Many seemingly innocent events can be transformed into crises. Western citizens have grown impatient with imperfections; they have come to fear glitches, and they see more of what they fear. In this culture of fear—sometimes referred to as the "risk society"—the modern mass media play an amplifying role (Ulrick Beck, 1992).

Challenges of Crisis Management

Crisis management has two dimensions. The technical-administrative dimension pertains to the coping capacity of governmental institutions and public policies in the face of emerging threats. But there is also a political dimension: Crisis management is a deeply controversial and intensely political activity. A combination of these dimensions translates into what Arjen Boin, Paul 't Hart, Eric Stern, and Bengt Sundelius have identified as the five critical challenges of crisis management: sense making, decision making, meaning making, terminating, and learning.

Sense Making

Crises seem to pose a straightforward challenge: Once a crisis becomes manifest, crisis managers must take measures to deal with its consequences. Reality is much more complex, however. Most crises do not materialize with a big bang. Policymakers must recognize from vague and contradictory signals that something out of the ordinary is developing. They must appraise the threat and decide what the crisis is about.

Crisis managers often find it hard to meet this challenge. The bewildering pace, ambiguity, and complexity of a crisis can easily overwhelm normal modes of situation assessment. Stress may further impair sense-making abilities. Organizational pathologies can produce additional barriers to crisis recognition.

Some categories of people are known for their ability to stay clearheaded under pressure. They have developed a mode of information processing that enables competent performance under crisis conditions. Veteran military officers, journalists, and fire and police commanders are known for this. Some organizations have developed a proactive culture of "looking for problems" in their environment. These so-called high-reliability organizations have somehow developed a capacity for thorough yet fast-paced information processing under stressful conditions. The unresolved question, as noted by Karl Weick and Kathleen Sutcliffe, is whether organizations can design these features into existing organizational cultures.

Making Critical Decisions

During a crisis, critical decisions must be made. Scarce resources may have to be prioritized. This is much like politics as usual, except that in crisis circumstances the disparities between demand and supply of public resources are much bigger, the situation remains unclear and volatile, and the time to think, consult, and gain acceptance for decisions is restricted. Crises confront governments and leaders with issues they do not face on a daily basis, for example, concerning the deployment of the military, the use of lethal force, or the radical restriction of civil liberties. Crisis decision making is making hard calls, which involve tough value trade-offs and major political risks.

An effective response to crisis requires coordination. After all, each decision must be implemented by a variety of organizations; effective implementation requires that these organizations work together. Getting public bureaucracies to adapt to crisis circumstances is a daunting—some say impossible—task. Most public organizations were originally designed to conduct routine business in accordance with values such as fairness, lawfulness, and efficiency. The management of crisis, however, requires flexibility, improvisation, redundancy, and the breaking of rules.

Coordination is not a self-evident feature of crisis management operations. The question of who is in charge typically arouses great passions. In disaster studies, the "battle of the Samaritans" is a well-documented phenomenon: Agencies representing different technologies of crisis management find it difficult to align their actions. Moreover, a crisis does not make the sensitivities and conflicts that governed the daily relations between authorities and others before the crisis disappear.

An effective crisis response is to a large extent the result of a naturally evolving process. It cannot be managed in a linear, step-by-step, and comprehensive fashion from a single crisis center. There are simply too many hurdles that separate a critical decision from its timely execution in the field.

Meaning Making

In a crisis, leaders are expected to reduce uncertainty and provide an authoritative account of what is going on, why it is happening, and what needs to be done. Leaders must get others to accept their definition of the situation. If they are not successful, their decisions may not be understood or accepted.

Public leaders are not the only ones trying to frame the crisis. Their messages coincide and compete with those of other parties, who hold other positions and interests and who are likely to espouse various alternative definitions of the situation and advocate different courses of action. If other actors succeed in dominating the meaning-making process, the ability of incumbent leaders to decide and maneuver is severely constrained.

It is often difficult for authorities to provide correct information right away. They struggle with the mountains of raw data (reports, rumors, and pictures) that are quickly amassed when something extraordinary happens. Turning these data into a coherent picture of the situation is a challenge. Getting it out to the public in the form of accurate, clear, and actionable information requires a major public relations effort. This effort is often hindered by the aroused state of the audience: People whose lives are deeply affected tend to be anxious and stressed (distressed). Moreover, they do not necessarily see the government as their ally.

Terminating a Crisis

Crisis termination is twofold. It is about shifting back from emergency to routine mode. This requires some form of downsizing of crisis operations. At the strategic level, it also requires rendering an account of what has happened and gaining acceptance for it. These two aspects of crisis termination are distinct, but in practice, they are often closely intertwined. The system of governance—its rules, its organizations, and its power holders—has to be (re)stabilized; it must regain the necessary legitimacy to perform its usual functions. Leaders generally cannot bring this about by unilateral decree, even if they possess the formal mandate to terminate crises in a legal sense. Formal termination gestures can follow but never lead the mood of a community. Premature closure may even backfire: Allegations of underestimation and cover-up are quick to emerge in an opinion climate that is still on edge.

Accountability debates can easily degenerate into blame games. Crisis leaders can be competent and conscientious, but that alone says little about how their performance will be evaluated when the crisis is over. Policymakers and agencies that failed to perform their duties prior to or during the critical stages may manage the crisis aftermath well, thus preventing losses to their reputation, autonomy, and resources. Crises have winners and losers. The political (and legal) dynamics of the accountability process determines which crisis actors end up where.

Learning

A crisis offers a reservoir of potential lessons for contingency planning and training for future crises. One would expect all those involved to study these lessons and feed them back into organizational practices, policies, and laws. This does not always happen, however. Lesson drawing is one of the most underdeveloped aspects of crisis management. In addition to cognitive and institutional barriers to learning, lesson drawing is constrained by the role of these lessons in determining the impact that crises have on a society.

The depiction of a crisis as a product of prevention and foresight failures would force people to rethink the assumptions on which preexisting policies and rule systems rested. Other stakeholders might seize on the lessons to advocate measures and policy reforms that incumbent leaders reject. Leaders thus have a large stake in steering the lesson-drawing process in the political and bureaucratic arenas.

Arjen Boin
Louisiana State University
Baton Rouge, Louisiana, United States

See also Leadership

Further Readings

Boin, R. A., ‘t Hart, P., Stern, E., & Sundelius, B. (2005). *The politics of crisis management: Public leadership under pressure.* Cambridge, UK: Cambridge University Press.

Janis, I. L. (1989). *Crucial decisions: Leadership in policymaking and crisis management*. New York: Free Press.

Perrow, C. (1999). *Normal accidents: Living with high-risk technologies* (2nd ed.). Princeton, NJ: Princeton University Press.

Turner, B. A. (1978). *Man-made disasters*. London: Wykeham.

Weick, K. E., & Sutcliffe, K. M. (2002). *Managing the unexpected: Assuring high performance in an age of complexity*. San Francisco: Jossey-Bass.

Critical Theory

Today, the term *critical theory* refers to a broad variety of theoretical approaches in social and cultural studies as well as social philosophy. Originally, critical theory was the programmatic name of a German group of philosophers and social scientists who, in 1930, began to work together under the directorship of Max Horkheimer at the Frankfurt Institute of Social Research and who continued their work after emigrating to the United States in the early 1930s. This group has also been labeled the "Frankfurt School" and is the subject of this entry.

From 1932 until 1941, the group published *Zeitschrift für Sozialforschung* (Journal for Social Research). After World War II, Horkheimer returned to Frankfurt and together with Theodor Adorno, who during the 1940s had become his most important theoretical companion, reestablished the Frankfurt Institute. Next to Adorno, among the important members of the first generation, are the philosophers Herbert Marcuse and Walter Benjamin, the economist Friedrich Pollock, the psychologist Erich Fromm, and the sociologist Leo Loewenthal.

With Jürgen Habermas, who in 1966 became Adorno's successor at the Frankfurt Institute, not only did a new generation take over, but a theoretical shift within the Frankfurt School appeared and soon became internationally recognized in political science and political theory in addition to philosophy. Nevertheless, Habermas in his social theory continued the critiques of "late capitalism" initiated by his predecessors. Today, some speak about a "third generation," referring to authors like Seyla Benhabib, Nancy Fraser, or Axel Honneth or focus primarily on Honneth as the successor to Habermas's chair and directorship of the Frankfurt Institute. However, the international diffusion and pluralization of the original ideas of critical theory renders questionable any claim to continuity of critical theory by intergenerational transmission. Compared with the first generation and Habermas, no such coherent school or theory exists any longer. The term *critical theory* today lacks a single specific meaning, and critical theory as a specific approach has, arguably, found its historical fulfillment with the life work of Habermas.

Origins and Early Development

The origin of critical theory has to be understood in the context of the deep political and theoretical crisis of the Marxist Second International during and after World War I, the Bolshevik Revolution, and the founding of the Soviet Union under V. I. Lenin. Perry Anderson has coined the term *Western Marxism* to summarize the different theoretical approaches of Georg Lukàcs, Karl Korsch, and Antonio Gramsci among others, who during the 1920s reacted to this crisis by creating their own rereadings of Marxism. In the late 1920s, this Marxist revisionism came to be strongly influenced by the first publication of Marx's early so-called humanistic writings. In all their variety, these new approaches were, on the one hand, opposed to the deterministic interpretation of history by "Marxism" (historical materialism) and, on the other, critical of the political instrumentalization of the Marxist approach as a mere power ideology in what soon came to be called "Marxism–Leninism."

When Horkheimer delivered his famous inaugural speech as director of the Institute in 1932, he introduced a research program that has been called "interdisciplinary materialism" (Wolfgang Bonß & Norbert Schindler, 1982). Its goal was to integrate, in a new manner, the Hegel–Marx tradition of philosophy of history with empirical social research in various disciplines into a new paradigm of practical philosophy. In this project, the role of philosophy was to integrate empirical research into a theoretical totality that could both represent reality and provide a normative perspective on it. Given the fact that in Western societies the proletariat had not proven to be the revolutionary class that would, according to Marx, overthrow the capitalist system of exploitation and alienation, the new paradigm of interdisciplinary materialism nevertheless was to remain devoted to the practical aim of emancipating society in totality by developing its inherent rational potential and its possibilities for liberation. The critique of ideology was one more traditional way to demonstrate the normative cleavages between the universal promises of the French Revolution—*liberté, égalité, fraternité*—and the unjust class realities of bourgeois society, based on a capitalist economy in which monopolistic structures had irreversibly undermined the socioeconomic foundations of liberalism. Despite the critique of liberalism, critical theory in its normative perspective always favored individualism over collectivism. Only an emancipated society could finally provide the social basis of "real individualism," while capitalism had reduced "individualism" to mere market-driven egoism and class antagonism.

In the face of the triumph of both Nazism and Stalinism, and with disappearing hope for a practical solution to the revolutionary question, Horkheimer reformulated the program of the group, in a famous article in 1937, as "critical theory." Critical theory, as opposed to "traditional theory" (i.e., positivism), should at least keep the consciousness alive in the absence of a revolutionary subject—that normative alternatives beyond the mere positivist affirmation of the given status quo in the fate of humankind were thinkable, at least in principle—that is, critical theory as the notorious "message in a bottle" in the absence of a practical subject of emancipation.

By the early 1940s, when Horkheimer and Adorno wrote their famous *Dialektik der Aufklärung* (*Dialectic of Enlightenment*; published in English in 1972), their diagnosis of time and history had become even more pessimistic. From the beginnings of enlightenment early in human history, the development of "reason" in the dominant form of "instrumental rationality" or "purposive rationality" alienated individuals and especially their system of social interaction from "objective reason" in constituting the subject—and thus their human potential. The history of humankind was thus—against Hegel's and Marx's views—perceived as, at best, an ambivalent story. In this pessimistic vision of mankind "logic of decay" at present, there was no longer a theoretical basis for hope of change. Later, in the 1960s, Adorno in his *Negative Dialectic* (published in English in 1973) confirmed this negative diagnosis in the topos of a "totally administered society." Marcuse, despite some political and personal distance from Horkheimer and Adorno since the late 1930s, came very close to this pessimistic approach with his concept of the one-dimensional man totally shaped by the main dominant values, in the early 1960s—though it was perceived quite differently by the international student protest movement of the late 1960s. The first generation of critical theorists, at least Horkheimer and Adorno, at the end of their theoretical path, which began without the practical subject of revolutionary practice, were deprived of a convincing theoretical foundation of rationality that could have been the background of their critical claim.

Habermas's Contributions to Critical Theory

It was Habermas who, as a former assistant to Adorno in the Frankfurt Institute, started a profound critique of the aporia concerning the way in which reason is held to be constituted. In Habermas's proclaimed "critical theory without praxis," he searched during the 1960s for a new immanent theoretical basis for the foundation of reason or rationality that could effectively claim to critique mere instrumental or purposive rationality. According to Habermas, the way back to Kantian transcendentalism would lead to normative philosophy only, which could not as such relate to the tradition of "materialist" social research. He critiqued the Hegelian naturalism inherent in the early program of critical theory; after some diversions, he began to develop a new foundation for critical theory in a linguistic turn away from the traditional Marxist approach. One could definitely ask the question whether traditional critical theory did not come to an end already with this fundamental shift in theoretical foundation, but Habermas vigorously, and finally successfully, always claimed the successorship.

Based on his reception of the work of Ludwig Wittgenstein, pragmatism, hermeneutics, and the theory of speech acts as developed by John L. Austin and John Searle, Habermas proclaimed that in the logic of communicative interaction a "telos of understanding" is always inherent. So he shifted the basis of socialization (*Vergesellschaftung*) from the Marxist emphasis on labor to the communication of people in their lifeworld. Properly analyzed and turned into reflexive practices, "communicative rationality" now could serve as the pragmatic *fundamentum* of various forms of rationality. This is the task of the *discurs* proposed and simulated at once in Habermas's new version of critical theory: In a discourse, the various implicit performative claims of participants become reflexive and, in Immanuel Kant's phrase, "under the coercion-free coercion of the better argument," accordingly transformative. Thus, in every real practice of discursive deliberation, participants have to agree on the normative principles of discursive justification.

Theory of Communicative Action

In his magnum opus, the *Theorie kommunikativen Handelns* (Theory of Communicative Action), Habermas elaborated the different forms of rationality, necessarily inherent if speakers claim the truth, rightness, or justice of what they say. In this analysis, he refers to a "counterfactual ideal speech situation," to which every person would have equal access, in which no power relations intervene and only the validity of arguments counts. According to Habermas, under such ideal conditions, participants would normally reach a *consensus* that, thus, could claim universal validity as reasonable. As one can see, at the heart of this argument lies the Kantian assumption that generalizability grounds normative universality.

The Public Sphere

While Habermas exposed this philosophical reconstruction of practical communication, he at once embedded it into a more empirically grounded social theory. Modern society—which he continued to describe as "late capitalist" (1975)—is functionally differentiated in "systems," such as the "political-administrative system" or the "economy," that are perpetuated through an abstract logic of purposive rationality, on the one hand, and the *Lebenswelt* (lifeworld), on the other, which is coordinated through communicative interaction; the "public sphere" functions in between, somehow intermediating. As an actual critical diagnosis, Habermas's thesis on a "colonization of the lifeworld by abstract systemic processes," especially its "economization," became very prominent and was popularized thereafter and perceived as the definitive new commentary of critical theory on the present society in "late capitalism." "Colonization" was substituted for the Marxist term of *alienation* in its critical function.

Critical theory always claimed to be more than just a normative theory by grounding the normative potential of its actual critique in the structures and logics of the society itself. While Habermas could claim that he has found this grounding on a microlevel in his discursive theory based on his pragmatic analysis of communication, for a critical theory of the present society, he had to demonstrate the applicability of his ideas on a macrolevel as well. Already, in his early *Der Strukturwandel der Öffentlichkeit* (*The Structural Transformation of the Public Sphere*), he reconstructed in an ideal-type manner the development and function of the *public sphere* (*Öffentlichkeit*) as part of the political and social emancipation of "bourgeois society" from feudal domination. In the institutions of this "public sphere"—journals, parties, associations, clubs, salons, and finally the parliament—emerging and prospering during the late 18th and the 19th centuries, the practice of discursive deliberation in public formed decisively the political and normative development of bourgeois society during its liberal period. *Publicity*—which had already played a role in Kant's philosophy—now emerged as a normative principle guiding the development of the rule of law (*Rechtsstaatlichkeit*) and the legitimization of political power. The "public sphere" thus represented in real history the societal requisite for public deliberation. Unfortunately and due foremost to the capitalist structure of the developing mass media, this public sphere, in the past liberal period of capitalism, had turned into an instrument of mass manipulation in the hands of big capital—leading to the "structural change" in the title of the book. So the normative functioning of the public sphere is just a *potential reality* that requires certain supporting conditions such as open and fair access for various social groups, freedom from censorship, and pluralistic media. In his later writings, Habermas in part also used the concept of *civil society* in which all sorts of informal groups, associations, and social movements contribute to this public process of normative and practical deliberation.

Deliberative Democracy

In recent years, Habermas has worked on reconstructing the interplay between the systemic reproduction of the economy (market) and the political-administrative system, on the one hand, and the norm-generating communication of lifeworld and civil society, on the other. In *Faktizität und Geltung* (Between Facts and Norms), Habermas offered a new account of the history of democracy and the rule of law (*Rechtsstaatlichkeit*) and applied it to the analysis of present Western democracies. Compared with the writings of critical theory up until the 1960s, Habermas now finally dropped the idea of any revolutionary transcendentalization of critical theory and reconstructed it as a normative critique based on the practical lifeworld perspective of participants. Against the imperatives of the purely purposive rationality of systemic reproduction in the fields of market economy and political administration, active participants ("citizens") in civil society involved in all sorts of mobilization, from protest movements to political parties, now have the chance to raise issues and to impregnate the systemic processes with the normative repertoire of lifeworld communication. Democracy, a specific version of it, now plays a central strategic role. Against the critique of mere liberal majority rule à la Joseph Schumpeter, Habermas develops his more philosophical discursive justification into a model of "democratic deliberation." According to this model, the democratic process does not just recognize and aggregate the prepolitical preferences of individuals as in the liberal model but views citizens as active participants in a process of public deliberation on practical and normative questions. This process is bound to what Philip Petitt (1992) has called "preference laundering": Participants are thus required to accept the principle of deliberative justification and change their original preferences in light of better arguments. Analogously to the normative philosophical model of discourse, the deliberative communications of the "public sphere" in the end also deserve the supposition that the generated outcome in the form of "public opinion" on a certain question is reasonable; a supposition, as Habermas continues to proclaim, that first depends on the approximation of the real communications process to the normative ideal requirements of discursive justification and, second, always remains fallible in the future.

In *Faktizität und Geltung*, this theory of deliberative democracy now serves as the normative foundation of the legitimacy of law; in its development, the system of law in a democratic society at best reflects the normative outcome of the lifeworld-based communications in practical and ethical issues in civil society; if and when the channels and institutions of the public sphere work properly, deliberation and communication are not distorted, and open access is given to the various normative preferences of a pluralist society.

The evolution of law under democratic conditions can thus in the long run be viewed as a social process of learning by which modern societies permanently adapt their system of positive law to new requirements in their internal and external environment. The "logic of decay," so characteristic of the main representatives of critical theory until the 1960s, has been gradually transformed by Habermas into an evolutionist framework of affirming the inherent potential of modern societies for normative and reasonable progress. In recent years, this more or less optimistic and somehow even neo-Hegelian view of historical progress has been expanded by Habermas to the evaluation of international law. While he vigorously critiques the unilateral interventionist policies of the United States under the George W. Bush administration, he has at the same time judged certain "humanistic interventions" based on moral grounds as "anticipating the universal rights of the people" and the reformed UN system as the core of a global legal system of justice.

Habermas's Influence in Political Science and Policy Studies

Habermas's philosophical works in recent years have had an enormous impact on analytical and theoretical developments within political science. In the political science literature, Claus Offe represents next to Habermas the turn to a more analytically based approach of critical theory; especially, the concepts of deliberation and justification through deliberation have widely spread from the subdisciplines of international relations to the empirical field of policy studies. In the former, Habermas's ideas have backed and reformulated the constructivist approach to international relations, such as that of Thomas Risse, and even inspired ideas of "cosmopolitan democracy," as in the work of David Held, while in various approaches to policy research, "deliberation in networks" is today viewed not only as a proper analytical tool but also as a practical device for more effective governance.

This cannot be covered here in detail. Only some questions and fundamental problems related to these adoptions of Habermas's normative theory can be briefly raised.

Most such problems arise in the process of transforming the "quasi-transcendental" normative theory of discourse and discursive justification of practical and ethical norms into a real model of policy making and democracy. The real political process in pluralistic societies has to take into account not only arguments but also particular interests and even beliefs and emotions that are not rationally grounded. Even the analytical distinction between processes of "arguing" for practical norms and "negotiating" for interests is not complex enough to cover all the *legitime* possibilities of expression of individual and collective preferences. The normative requirement of participation in public deliberation—namely, of preference transformation in discourses—can turn into illiberal coercion against individuals and groups, which might deprive them of their privacy.

As far as the future of "democracy beyond the state" (see Michael Greven & Louis Pauly's 2000 book by that title) is concerned, another problem related to the deliberative conception of democracy becomes even more evident and, indeed, already is on the level of nation-states: Who can and who should participate in deliberative governance networks as they have developed and are proposed on the transnational level? Not every kind of participation in arrangements of "participatory governance" is democratic, because often, the essential normative requirement insisting on the equality of participants is violated by the "participation" of collective actors and all kinds of organizations. The normative principal–agent problem refers to collective actors who also claim to be advocates of public interests.

Confronted with the ongoing and seemingly irreversible process of denationalization and growing interdependence of former state-organized democracies, which both threaten the idea of congruence between a democratic demos and a territorial-based regime of governance, critical theory should perhaps return to its original historical approach and start to think about a future beyond traditional democracy.

Michael Th. Greven
Universität Hamburg
Hamburg, Germany

See also Critical Theory in International Relations; Kant, Immanuel; Liberalism; Marxism; Positivism

Further Readings

Adorno, T. W. (1973). *Negative dialectics.* New York: Seabury.

Bonß, W., & Schindler, N. (1982). Kritische Theorie als interdisziplinarer Materialismus [Critical theory as interdisciplinary materialism]. In W. Bonß & A. Honneth (Eds.), *Sozialforschung als Kritik. Zum sozialwissenschaftlichen Potential der Kritischen Theorie* [Social research as a criticism. To the sociological potential of the critical theory]. Frankfurt, Germany: Suhrkamp.

Habermas, J. (1975). *Legitimation crisis.* Boston: Beacon Press.

Habermas, J. (1984–1987). *The theory of communicative action* (Vols. 1–2). Boston: Beacon Press.

Habermas, J. (1989). *The structural transformation of the public sphere.* Cambridge: MIT Press.

Habermas, J. (1996). *Between facts and norms.* Cambridge: MIT Press.

Horkheimer, M. (1982). *Critical theory.* New York: Seabury.

Horkheimer, M., & Adorno, T. W. (1972). *Dialectic of enlightenment.* New York: Seabury.

Marcuse, H. (1964). *One-dimensional man.* Boston: Beacon Press.

Critical Theory in International Relations

The development of explicitly critical approaches to the study of international relations (IR)—that is, critical IR theory—is generally dated to the early 1980s. Today, it can be seen as a pluralistic family of theories, containing both modernist and post-modernist variants, including Frankfurt School–influenced IR theory, Gramscian international political economy, feminist IR theory, postmodern IR theory, and most recently, postcolonial theory.

What these admittedly disparate traditions have in common is a commitment to the enterprise of critique. At its most general level, this can be understood in terms of the distinction, common to the work of many critical IR theorists, between two forms of theorizing—"problem solving" (or "traditional"), on the one hand, and "critical," on the other. Proponents of critical theory start from the proposition that theory is never neutral but always constructed in response to a specific agenda and in the service of a specific project. Problem-solving and traditional theory—of which mainstream IR theory stands as a prime example—is judged by the proponents of critical theory to be a form of theorizing that stands in the service of the status quo. The knowledge associated with mainstream IR theory is in their view not value free; rather, it is generated to help society's elites maintain their positions of power and privilege, and, at its most extreme, it is in the service of a totalitarian project on a global scale.

In contrast, critical forms of theorizing are characterized by the following elements.

The first element common to these critical forms of theorizing is methodological opposition to positivism in mainstream IR theory. From the perspective of critical theory, positivism, with its defining characteristics of naturalism, the separation of fact and value, and the separation of the observed and the observer, is more than just an approach to the study of social phenomena. Rather, it contains within it a telos of domination—a telos that derives from positivism's interest in constituting the world for the purposes of control and from its tendency to reinforce the status quo by treating the socially constructed as given and by representing that which is alterable as something that cannot be changed.

It should be noted that the critique of the highly normative agenda of mainstream IR theory is not an argument about the ideological biases of mainstream theorists and how they inform and distort the social-scientific enterprise. Such biases may well exist, but in the view of critical theorists, it is the knowledge-constituting interests of mainstream IR theory itself, not the values of individual researchers, that give it its conservative politico-normative agenda. Therefore, mainstream researchers cannot overcome their conservative bias by being more vigilant in their research design or by being more self-aware of their biases.

In contrast, critical theory integrates so-called qualitative methodologies aimed at unearthing the way intersubjective meanings are both embedded in and instantiate social practices. The objective here is not to deny the existence of regularities in the social world but rather to determine when such observable regularities are regularities *tout court* and when they can be understood to be the manifestations of ideologically petrified relationships of domination that could be challenged and overcome.

A second element of critical forms of theorizing is the adoption of a strategy of immanent—as opposed to transcendent—critique. Where the latter applies transhistorical (and therefore, ahistorical) standards to objects or phenomena under examination, immanent critique applies historically emergent standards for the purposes of criticism that are generated by the object/phenomenon itself.

A third element is a shared commitment to promoting self-consciously normative discussions within the discipline. What is distinctive about this activity, moreover, is that in light of the critique of the separation of fact and value noted above, critical theory's promotion of normative critique is formulated so as not to reproduce the problematic separation of fact and value characteristic of positivism. That is, critical IR theory problematizes the conventional distinction between "normative" and "empirical" theory, noting the inherently normative content of all forms of theorizing, including so-called "empirical" efforts.

A fourth element is the critique of ideology. From the conclusion that neorealism and neoliberalism are two sides of the same mainstream coin, to the critique of the gendered nature of the global political economy (not to mention the theories developed to "explain" it), and to the way dominant academic discourse continues to contribute to the view that the underdeveloped world is a simple casualty of inherent weaknesses and limitations, critical IR theory carries its agenda of ideology critique forward on a number of different levels and in terms of a range of distinct, but interrelated, foci (class, race, gender, sexual orientation, etc.).

A fifth element is a shared interest in questions of oppositional practice and their potential to contribute to emancipatory change. Here, the focus is on social forces—social movements, old and new (e.g., labor, ecological movements, women's movements, and human rights)—and the way in which critical theory can help explain the emergence and nature of these movements as well as how critical IR theory can generate knowledge that might prove to be of use to these movements.

As already noted, differences exist between critical IR's subtraditions. An early tension was one between modernist—enlightenment-inspired forms of theorizing—and postmodernist challengers. The latter's skepticism about metanarratives in general—including metanarratives framed in terms of the telos of human emancipation—raised fundamental questions both about the viability of modernist variants of critical IR theory and about the political content of postmodernist variants.

Beyond this tension, others have manifested themselves. An early success by mainly postmodernist critical theorists in being given a special issue of the *International Studies Quarterly* was marred by their failure to include even one contribution from a feminist theorist. Wariness on the part of the feminist IR community with regard to "masculinist" theorizing—including critical theorizing—continues (and not without justification). More recently, postcolonial theorists (who include a number of feminists) have suggested that insensitivity to matters of race as well as gender leaves at least some critical IR theory open to the charge that it is not only masculinist but "White supremacist" as well.

In recent years, a number of efforts have been made to assess the impact of the critical theoretical tradition on the discipline as a whole. Some have argued that the critical engagement with positivism and the openness to interpretive methodologies have helped create greater "thinking space" within the mainstream IR theory (see, e.g., the movement within mainstream IR known as constructivism). This judgment gives rise to the following question: Having had a positive influence on mainstream IR theory, does an explicitly oppositional critical theory have any further use in the discipline?

More skeptical critics have expressed doubts about whether the metatheoretical discourse of critical IR theory has done more than waste intellectual energies that might have been better put to use studying the concrete problems of world politics. Some have even gone so far as to charge critical theory with an ideologically motivated "cleansing" agenda, with mainstream approaches such as realism being driven increasingly to the margins of the discipline.

In addition, some who profess to share ongoing sympathies with the critical project nonetheless expressed doubts about its adequacy in terms of practice—specifically, critical IR's failure to make meaningful links to antisystemic movements—whether as a result of inadequacy at the level of conceptualization or in terms of the origins and prejudices of critical theorists themselves.

Finally, the origins and geographical presence of critical IR theory itself is worth noting. The discipline of IR is overwhelmingly a discipline of the English-speaking world, with the United States being the unquestioned center in which the bulk of IR theory is produced. Significantly, critical IR theory is largely a product of the academic "semiperiphery." That is, while some critical IR theorists are to be found within the borders of the United States, a disproportionate number are products of and are academically active in the United Kingdom, Canada, Australia, and New Zealand. A sociology of knowledge account of this aspect of critical IR theory remains to be written.

Mark A. Neufeld
Trent University
Peterborough, Ontario, Canada

See also Feminist Theory in International Relations; Hegemony; Positivism; Postcolonialism; Postmodernism in International Relations; Realism in International Relations

Further Readings

Ashley, R. A. (1984). The poverty of neorealism. *International Organization, 38*(2), 225–286.

Cox, R. W. (1994). *Approaches to world order* (T. Shaw, Ed.). Cambridge, UK: Cambridge University Press.

Ling, L. H. M. (2001). *Postcolonial international relations: Conquest and desire between Asia and the West.* Basingstoke, UK: Palgrave Macmillan.

Neufeld, M. (1995). *The restructuring of international relations theory*. Cambridge, UK: Cambridge University Press.
Review of International Studies, 33 (S1). (2007). [Special issue on critical IR].

Cross-National Surveys

Comparative research has been broadly applied by the social sciences research community for at least 60 years now. Its application in disciplines such as cultural or social anthropology counts among the most important advancements in the contemporary understanding of cultures and societies. In political science, cross-cultural studies have increasingly been based on cross-national surveys. The word *survey* itself is broad, even if many researchers use it with a "quantitative" meaning: the collection, through a questionnaire survey and on a representative sample, of observations coming from the quantitative coding of the responses chosen by respondents. In the following, the use of the word *survey* will follow this tradition. This entry discusses the developments in this field and addresses remaining problems.

Evolution of Cross-National Research

The developments of European integration and the interdependences created by globalization have clearly increased the value of comparisons across different national systems and across cultures; we are in the context of what Anthony Heath calls the "globalization of public opinion research." This is reflected in the greater number of available sociopolitical "barometers." Following the example of the Eurobarometer since the 1970s, these are now combined in the "Global Barometers" project, including also Afrobarometer, Latinobarómetro, and East Asia Barometer.

Despite these significant recent developments, cross-national survey methodology has a long history. The process started more than 4 decades ago, with Gabriel Almond's and Sidney Verba's seminal contribution in *The Civic Culture* (1963), which was updated and revised in 1980. Previously, a few other cross-national surveys of political attitudes had been developed, notably William Buchanan and Hadley Cantril's *How Nations See Each Other* (1953), sponsored by UNESCO (United Nations Educational, Scientific and Cultural Organization). In a recent overview of available cross-national surveys, the German data archive (Zentralarchiv in Cologne) found that from 1948 (fieldwork date for *How Nations See Each Other*) to 2004, more than 60 projects of cross-national surveys have taken place. Profusion of data rather than a lack of data is, in fact, the case today. But profusion does not necessarily mean availability of "truly" comparative, harmonized, and replicated studies. In the beginning of the 1990s, a European Science Foundation (ESF) paper reported that Europe was rich in social sciences data but poor in harmonized and comparable data. Since that period, much progress has been made (in particular due to the European Social Survey, ESS), but a few important methodological problems comparing surveys remain.

The profusion of available data makes it impossible to present an exhaustive picture of cross-national surveys. It is more fruitful to present some of the major series and the methodological issues that they pose in terms of measurement, comparison, and data analysis. In another recent survey of major comparative studies, Pippa Norris (2009) presents a quasi-exhaustive list of the most relevant ones. She proposes to keep as most relevant the studies "defined as those covering more than one independent nation-state which have established a regular series of surveys of social and political attitudes and behavior" (p. 522). This includes the Eurobarometer and related European Union (EU) surveys (which started in 1970), the European Elections Study (1979), the European Values Survey and the World Values Survey (1981), the International Social Survey Programme (1985), the Global Barometers (from 1990), the Comparative National Elections Project (1990), the European Voter and the Comparative Study of Electoral Systems, the ESS (2002), the Transatlantic Trends survey (2002), the Pew Global Attitudes project (2002), and the Gallup World Poll (2005). In the following, only a few of these studies can be discussed in greater detail.

The Eurobarometer

Among these, the Eurobarometer is the most important with regard to its time series and its coverage. The Eurobarometers, initiated, developed, and organized by the European Commission, have played a key role in the expansion of political and social attitudinal measurement across Europe. Initiated by Jacques-René Rabier, a high-level civil servant at the European Commission having visionary ideas about the role of citizens in the construction of a legitimate Europe, this series investigates mainly the key question of European citizens' attitudes toward European integration. The series has been, and continues to be, more than just a public communication tool for the EU Commission. Its findings and data have been extensively used by social scientists in particular in two fields of political sociology: the analysis of attitudes toward EU (integration, enlargement, EU domains of public policies, citizenship and national attachment, trust and knowledge about EU institutions, and support for EU) and the analysis of the postmaterialist "revolution." This research, initiated by Ronald Inglehart, was clearly linked, at least in the first steps, to Eurobarometer survey items.

Even if academic researchers have many opportunities to use Eurobarometer data, this project has its own agenda and is largely, if not exclusively, directed toward EU matters. Conducted two times a year since 1973, the "regular" Eurobarometer series (the EU Commission also organizes other "special" Eurobarometer studies on focused topics) offers a unique time series to investigate trends in support for the EU. A recent development of the questionnaire concerns the introduction of items related to globalization and the attitudinal connections between perceptions of the EU and perceptions of globalization. The Eurobarometer covers the population of the respective nationalities of each EU member state aged 15 years and older. An interesting point of the series is that it also samples the populations of candidate countries (in 2010, Croatia, Turkey, and the former Yugoslav Republic of Macedonia; it also samples the Turkish Cypriot Community). With few exceptions (e.g., Luxembourg), the sample sizes of national samples are about 1,000.

European Values Studies

A second well-established cross-national survey series is the European Values Studies (EVS), from which the World Values Study (WVS) developed. Funded by a private foundation (the EVS foundation, based at Tilburg University in the Netherlands), the primary focus of the EVS is on changing moral and social values, with particular emphasis on Christian values and their possible replacement by alternative "values systems" (see Loek Halman, 2001). The EVS series is profoundly marked by this original intellectual project, whose goal is to understand the value changes across Europe, being investigated in specific domains such as family, religion, morality, work, and politics. Some new issues were covered in the last EVS wave, such as environmental concerns. The questionnaire of the EVS series is among the most extensive in contemporary social surveys, and the contribution of the EVS series to comparative political sociology is certainly very important; the quality of the questionnaire and its methodological and substantive integration are of a very high standard. A major difference with the Eurobarometer series is the restriction of the EVS questionnaire to a set of articulated and intellectually integrated research questions and items. The first round of the EVS was conducted in 1981, the second in 1990, the third in 1999/2000, and the last one in 2008/2009. EVS is completely social science driven, managed by Tilburg University, in connection with national academic teams. It is conducted only every 8 to 9 years because political and social values do not change rapidly. In other words, the EVS perspective is on gradual cultural changes rather than on the yearly monitoring of a public opinion barometer. The exclusive academic orientation of the EVS series also makes a big difference compared with the Eurobarometer: A "theory group" and a "methodology group" investigate and monitor closely the production of the questionnaire and the harmonization of fieldwork, translations, and data coding protocols. The last wave of the EVS (2008/2009) has in particular been the object of intensive methodological analysis by these two committees. The EVS is now composed of 46 national representative samples (there were 16 countries in the EVS series in 1981, 23 in 1990, and 33 in 1999). The sampling procedure for EVS is, for each country, a representative (multistage or stratified random probability) sample of the adult population of the country aged 18 years and older. The sample size is about 1,500 respondents per country. Although EVS is rather different from Eurobarometer, it is similar to it in that it is not a "general social survey."

The International Social Survey Programme

The International Social Survey Programme series (ISSP) is also not a general social survey, but it is academically oriented and managed like the EVS. It is conducted on a yearly basis, which makes it possible to follow a set of indicators over time. From the perspective of a questionnaire design, the particularity of ISSP is the "rotating-modules" component: Every year the questionnaire is devoted to a particular topic (e.g., religion, role of government, and social inequalities). Each module is designed to be repeated at nonregular intervals to allow measures of change over time and between countries. From the population coverage point of view, ISSP has an international perspective with a broad and diverse set of countries: In 2010, 44 countries from all continents participated in ISSP, even if South Africa remained the only African country. The methodological rules of ISSP (the harmonization of sampling designs, questionnaire translation, and modes of interview) are quite strict. ISSP is organized with an annual general assembly that decides key issues, in particular the choice of the rotating modules. The lack of resources for permanent management weakens the strong methodological requirements somewhat, and ISSP is more like a collection of highly integrated and harmonized national studies than a truly pan-international study. Samples are representative samples of the population of the countries aged 18 years and older; no telephone interviews are permitted (but alternatives to face-to-face interviews are possible, e.g., postal surveys); and sample sizes and sampling designs may vary somewhat from country to country.

The European Social Survey

The ESS is the most recent cross-national data enterprise and, surely, with EVS, the most advanced one in terms of methodological controls. The ESS found its original motivation in a "blueprint" document coming from a group of experts commissioned by the ESF, following the ESF program "Beliefs in Government." The ESS has many particularities in funding (it has been funded through the European Commission's Framework Programmes, the European Science Foundation, and national funding bodies in each country), in organization (the ESS is directed by a Central Coordinating Team, CCT), and in methodological excellence. The ESS can be considered as one of the most integrated and harmonized comparative projects in the social sciences, and in 2005, it has been awarded Europe's very prestigious annual social science award—the Descartes Prize. Five waves (the fieldwork for the fifth round finished in early 2011) have been completed on random samples of resident populations over the age of 15 in most EU countries as well as in some non-EU member states (e.g., Switzerland and Norway). Depending on the waves, a set of about 30 countries is covered by ESS, which runs every 2 years. The technical specification of the ESS requires that a minimum "effective achieved sample size" should be 1,500 respondents, or 800 in countries with populations of fewer than 2 million.

The ESS represents a very significant step forward in the methodological controls before, during, and after the fieldwork period. The methodological high standards have been a key issue for the ESS from its beginning, and the establishment of a permanent team taking decisions, supporting, and controlling strictly the national teams is an important aspect of this organization. To get the official label of "ESS," a national team has to strictly follow and implement a methodological and organizational template, which covers sampling (strict random sampling), translation of the questionnaire, fieldwork administration, documenting, and delivering the different data files to the NSD archive in Bergen, Norway. Since the creation of the ESS in 2002, many scholars across the world have used it, and very significant improvements of survey practices and data uses have been made. The CCT is supported by a number of advisory and consultative groups, principal among which is the Scientific Advisory Board under the chairmanship of Max Kaase. The Scientific Advisory Board consists of one representative each from the participating national funding agencies, one from the ESF, and two from the Commission. A small multinational "methods group" advises on other methodological issues. For every round of the ESS, a core set of questions is replicated (achieving in so doing the objective of a regular general social survey), while two "rotating modules" are chosen after an open and international competition between research teams.

Methodological and Substantive Issues

This short and selective overview of cross-national surveys leads to a set of conclusions. Very significant progress has been made on the data quality and on the harmonization of sampling designs, questionnaire translations, equivalence of measures, and data analysis. The problem of functional equivalence persists, however, and using cross-national surveys requires a lot of attention to methodological concerns. The long and sometimes difficult road to the production of truly comparative surveys is certainly not yet at its end, even if the ESS has achieved a high standard of methodological controls. Methodological and substantive issues are closely connected in achieving comparative cross-national data; comparative analysis is (or should be) about testing if apparent commonalities or differences are "true" and to investigate the role of deep (thick) idiosyncrasies (e.g., of history, institutions, and territories) in those findings. Even with the best intentions (as in the studies mentioned above), some methodological problems remain. For example, response rates have some variance across countries, difficulties in translating some items and concepts persist, and controlling for fieldwork specificities is critical. Documenting and reporting about these issues as well as details about interviewers and fieldwork is very important. Knowing the possible sources of errors and misinterpretations in the comparative analysis of cross-national surveys is, indeed, fundamental. The achievements of the different data archives (e.g., Council of European Social Science Data [CESSDA] archives, Interuniversity Consortium for Political and Social Research [ICPSR], or comparative data labs such as European Centre for Analysis in the Social Sciences [ECASS] or Network of Economic and Social Science Infrastructures in Europe [NESSIE]) in this respect are already remarkable but not yet finished. In a digital world, the easy access of users to very diverse collections of data, but those that are sometimes not comparable and not fully documented, creates an even greater need for infrastructures with permanent duties and resources for harmonizing, documenting, archiving, and disseminating data and training the users.

The profusion, diversity, and availability of cross-national surveys also raises a set of new methodological issues. In increasing the *N* (in both space and time), the availability of new data sets has returned an old question on the agenda of comparativists: the challenging intellectual project defined in 1970 by Adam Przeworski and Henry Teune to replace the proper names of nations (or states or cultures) with the names of variables. Cross-national surveys, in particular when they are conducted in a large number of countries and/or repeated through time on a regular basis, make it possible to think about this variable-oriented approach. If this project is challenged by case-oriented approaches, because cultures, nations, or states are not just "control" variables, the availability of cross-national surveys such as the ESS, EVS, ISSP, or Eurobarometer at least makes it possible to go in that direction. The return of this old question also challenges the statistical techniques used to analyze the data, in particular to test for country effects. Many users still simply juxtapose their results country by country, not really facing the logic of statistical control for country effects, or they just introduce countries as dummy variables in their models. But this is not enough. Recent advancements in statistical techniques make it possible to analyze the data coming from studies such as the ESS, EVS, ISSP, or Eurobarometers with modeling techniques with more sophisticated contextual analysis techniques (e.g., multilevel models, latent class analysis, and multivariate techniques coping with country effects). A new challenge is thus going to be the production of data sets articulating the microdata (respondents) with proper contextual variables not reduced to the macrovariables at the national level. Availability of contextual variables at the subnational level will be an important concern as well as the availability of political contextual variables "closer" to the respondents than the national ones. The impact of institutional designs on political attitudes and behaviors also could benefit greatly from more detailed contextual data.

Bruno Cautrès
CEVIPOF/Sciences Po
Paris, France

See also Attitudes, Political; Beliefs; Election Research; Political Culture; Postmaterialism; Survey Research; Survey Research Modes; Values

Further Readings

Almond, G., & Verba, S. (1965). *The civic culture*. Boston: Little, Brown.

Almond, G., & Verba, S. (1980). *The civic culture revisited*. Boston: Little, Brown.

Buchanan, W., & Cantril, H. (1953). *How nations see each other*. Urbana: University of Illinois Press.

Halman, L. (2001). *The European values study: A third wave* (Sourcebook of the 1999/2000 European Values Study Surveys). Tilburg, Netherlands: EVS/WORC/ Tilburg University.

Heath, A., Fisher, S., & Smith, S. (2005). The globalization of public opinion research. *Annual Review of Political Science, 8,* 297–333.

Jowell, R., Roberts, C., Fitzgerald, R., & Eva, G. (Eds.). (2008). *Measuring attitudes cross-nationally: Lessons from the European social survey*. London: Sage.

Norris, P. (2009). The globalization of comparative public opinion research. In T. Landman & N. Robinson (Eds.), *The SAGE handbook of comparative politics* (pp. 522–540). London: Sage.

Cross-Tabular Analysis

Cross-tabulation is the statistical technique by which two or more discrete random variables (usually variables that can take only on a finite number of distinct values) are cross-classified to fully characterize their joint distribution and thereby the relations between them. The analysis of the resulting cross-classification of frequencies, or contingency table, is most often carried out to establish relations of (in)dependence between the variables and to gauge the strength of these relations, although several other objectives are equally well served with the use of cross-tabular analysis. This type of statistical analysis is relatively simple to understand, easy to implement, and well suited for a variety of interesting questions. As follows, its major features using a simple example from political science are discussed.

To illustrate the use of cross-tabular analysis, consider the following hypothetical example. A researcher is interested in understanding the cosponsorship decisions of legislators. From a random sample of legislators, she or he has collected data on the number of times legislators cosponsored with members of the parties they ran against in the previous election (i.e., their adversary parties) as well as a dichotomous measure of how electorally vulnerable legislators reported to have felt in the previous election. She or he is interested in knowing (a) whether these two variables are discernibly related and, if they are, (b) how sizable the effects of the relation are.

A two-way (i.e., including two variables), two-by-two (each taking on two distinct values) contingency table reports the results of her or his data collection (Table 1). In this case, variables C (for *C*olumn and *C*osponsorship) and R (for *R*ow and *R*eported vulnerability) were laid out in this particular order following the convention of placing the categories of the response variable in the columns and the categories of the explanatory variable in the rows. This is mere convention, and transposing the table would not lead to different conclusions. Three sample distributions can be derived that fully summarize the component variables: (1) the sample joint distribution of R and C, (2) the sample marginal distribution of R (or C), and (3) the sample conditional distribution of R given C (or C given R).

The joint distribution of R and C gives the probability of classifying an observation in a particular cell of the table and can be approximated by calculating the proportion of the data that fall into each cell; this is the sample joint distribution. The marginal distribution of R (or C) is the probability of classifying observations in a particular row (or column) of the table, and once more, it can be approximated by the sample marginal distribution of R (or C), obtained by summing across columns (or rows) and dividing by the total amount of observations. Finally, the conditional distribution of R (or C) comprises the probabilities of classifying an observation in a particular row (or column) given that they have been classified in a particular column (or row), and it can be approximated by dividing the proportion of observations being classified in a particular cell over the overall proportion of observations being classified in the corresponding column (or row), yielding the sample conditional probability of R given C (or C given R).

To answer question (a) from the previous page, the researcher must evaluate the proposition of independence. In the cross-tabular setting, independence means that the probability of classifying

Table 1 Contingency Table of Cosponsorship Patterns Versus Reported Vulnerability

	Bills Cosponsored With Adversary	*Bills Not Cosponsored With Adversary*
Reported vulnerability	568	287
Did not report vulnerability	94	246

an observation in a particular cell is equal to the product of the (marginal) probability of classifying an observation in the corresponding row and the (marginal) probability of classifying an observation in the corresponding column. The difference between the values expected under independence and the values actually observed becomes the basis for evaluating question (a). More formally, the researcher must calculate

$$\text{Pearson } \chi^2 = \sum_{r=1}^{R} \sum_{c=1}^{C} \frac{(n_{rc} - e_{rc})^2}{e_{rc}},$$

where n_{rc} is the observed count at cell rc, and e_{rc} is the expected count at cell rc under independence (i.e., the product of the marginals). The researcher then compares this test statistic with an X^2 distribution with $(R - 1)(C - 1)$ degrees of freedom. First, the test of question (a) (i.e., the test of independence) can be carried out by obtaining the test statistic, choosing a significance level for the test α, and then comparing the value of the test statistic with that of the quartile of the X^2 table for $(R - 1)(C - 1)$ degrees of freedom and the chosen α level. If the value of the test statistic is higher than the resulting quartile obtained from the table, the researcher rejects the null hypothesis of independence.

This procedure is known as the Pearson X^2 test. It is, by far, the most widely used test of independence and, for a large number of observations, is equivalent to other tests for cross-tabulated data—namely, the randomization X^2 and the likelihood ratio X^2. As a rule of thumb, the expected cell frequencies under the null hypothesis of independence should exceed 5 for all cells for the test to produce reliable results. When cell sparseness is an issue, Fisher's exact test of independence is recommended instead. Also, if at least one of the component variables is ordinal, more powerful (i.e., more likely to reject a false null hypothesis) tests are available, and therefore, their use is recommended whenever appropriate (e.g., the mean score statistic test).

In the above hypothetical case, the researcher finds a test statistic of 148.1, which was not likely drawn from an X^2 distribution with $1 = (R - 1)(C - 1)$ degrees of freedom. Hence, the evidence supports the idea that these two variables are, in fact, not independent.

Next, the researcher will try to establish how big a substantive effect this dependence entails—question (b). The most commonly used measure of effect magnitude is the odds ratio. Odds are a measure of relative probabilities, which convey how much more likely an outcome is than its alternative. Hence, odds ratios are themselves relative measures of how much more likely an outcome is than its alternative for a given individual than it is for another one, defining individuals as actors with different probabilities assigned to each outcome. These ratios can take any nonnegative value and are usually expected to be different from one—a ratio of one implying no difference in the odds of an event for the two different individuals (or compared groups). Therefore, they can be used to investigate the effects of belonging to different categories of the R variable on the odds of belonging to a particular category in the C variable. To calculate odds ratios, the sample joint probability distribution is needed again, and the proportions of observations of the cells are denoted as p_{rc}. In general, the odds ratio in a two-by-two setting is calculated by

$$\Omega = \frac{p_{11}p_{22}}{p_{21}p_{12}}.$$

The fact that main and off-diagonal elements are multiplied and divided earned the odds ratio the name "cross-product" ratio, and this is also the reason why the results are impervious to table transposition. In the case of our researcher, the odds of vulnerable legislators sponsoring a bill with their adversaries are estimated to be 5.18 times the odds of a safe legislator sponsoring a bill with adversaries. Finally, a confidence interval can be calculated for this estimate of the odds ratio based on the normal distribution. By taking the natural logarithm of the odds ratio, its variance is defined as

$$\mathrm{Var}[\log(\Omega)] = \frac{1}{n_{11}}\frac{1}{n_{12}}\frac{1}{n_{13}}\frac{1}{n_{14}},$$

making the expression for the $(1 - \alpha)$% confidence interval

$$\exp\left(\log(\Omega) \pm z\sqrt{\mathrm{Var}(\log(\Omega))}\right),$$

where z is the standard normal quartile at the $(1 - \alpha)$ cumulative probability. In the case of our researcher, the 90% confidence interval around the estimated odds ratio is (4.11, 6.53).

Once again, if one of the random variables in the cross-tabulation is ordinal, other more powerful tests are available—namely, the gamma test, Kendall's tau *b*, and the lambda test (if *both* variables are ordinal).

These methods can be extended to situations involving more than two rows and more than two columns, and they can even be applied when other covariates are expected to mediate the relation between a pair of random variables. In these cases, researchers are interested in partial associations between random variables, conditional on the values of any number of "control" or "stratifying" variables. Failure to control for these variables can result in spurious results. In practice, for cross-tabulations with more than three variables, these methods become increasingly harder to perform, and more sophisticated estimation procedures such as those involved in fitting logistic models are recommended.

Finally, although these are the two main uses of cross-tabulation, there are other questions that can be addressed via this technique. In particular, it is possible to perform cross-tabular analysis in order to judge homogeneity across populations, to answer questions about differences in proportions across groups, or to judge agreement between, for instance, observers evaluating the same phenomenon. In the last case, the preferred statistic is called Cohen's kappa.

Santiago Olivella
Washington University in St. Louis
St. Louis, Missouri, United States

See also Categorical Response Data; Logit and Probit Analyses; Quantitative Methods, Basic Assumptions; Statistical Inference, Classical and Bayesian

Further Readings

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). New York: Wiley.

DeGroot, M. H., & Schervish, M. J. (2002). *Probability and statistics* (3rd ed.). New York: Addison-Wesley.

King, G. (1989). Models for tabular data. In *Unifying political methodology: The likelihood theory of statistical inference* (pp. 133–161). Cambridge, UK: Cambridge University Press.

Rudas, T. (1998). *Odds ratios in the analysis of contingency tables*. Thousand Oaks, CA: Sage.

Swafford, M. (1980). Three parametric techniques for contingency table analysis: A nontechnical commentary. *American Sociological Review, 45,* 664–669.

CULTURALISM

Culturalism, which would ground explanations of political phenomena in cultures rather than in notions of structure or agency, is problematic because of the lack of consensus on how to define culture. Moreover, there is little agreement on the added value of culturalism: Whether culture is the main explanatory variable or a residual intervening factor is a cause of enduring debate among scholars. Culturalism may be seen as belonging to the archaeology of politics as do structuralism, functionalism, and behaviorism, once very popular themselves. Alternatively, it could be viewed as living a new life, rejuvenated within the field of "cultural studies," which is nevertheless little concerned with politics but tries to capture all sorts of discourses and ways of life. Finally, it may be overlapping with emerging paradigms that make an extensive use of culture without claiming it—such as constructivism and cognitive science.

Discussing culturalism has some virtue. First, it brings to the fore some particular aspects of social and international life that are ignored by "structural" or "rationalist" theories—such as identity issues, which are still divisive everywhere. Second, its anthropological stance enriches overly parsimonious analyses of decision making that rely on a limited set of factors to fit scientific standards, as does "rational choice." Third, whereas its birth coincides with the demise of functionalism—whose self-proclaimed biological inspiration from the 19th century onward was a source of numerous caveats—culturalism was at its apex when political economy failed to explain change. At the very least, then, it was instrumental in discrediting "organicism" and Marxism. Fourth, although excluded from most comparative research, it recently made a great comeback in international studies, where the topics of conflicts between civilizations and intercultural negotiations have become popular. Therefore, it is now taken for granted that culturalism has some merits to be properly evaluated.

Unfortunately, culturalism itself cannot be easily defined, since culture can be

- what makes a polity distinct ("Culture" with a capital C);
- a set of shared beliefs about causation, norms, and values within a local, domestic, or international community ("culture" in the singular); or
- shorthand for a long list of things that people living in the same country have in common, from lifestyles to sources of national pride, not to forget a common understanding of what alternative policy issues are ("cultures" in the plural).

As for the term *culturalism*, it suggests

- awareness of the possible contribution of cultural variables to the aggregated explanation of a political phenomenon ("thin" culturalism) or
- privileging cultural factors over more political variables such as social stratification, ideological and partisan cleavages, political attitudes, institutional constraints, and strategic choice ("thick" culturalism).

Because science is deterministic and multicausal, the second and fourth definitions are adopted here. That both lack precision stems from the ontological status of culture: Becoming visible only through transmission from one generation or country to the next (i.e., the "socialization" and "acculturation" processes), it can only be deducted from its impact on receivers.

A Short Genealogy

Until the end of the Cold War, sociologism and functionalism were deeply opposed in methodology as well as with respect to the goals of social science. On the left, sociologists of all creeds predicted the end of capitalism, exploitation, alienation, and imperialism. On the right, functionalists assumed that social change was not as effective as expected: Because basic functions had to be fulfilled at any rate, a "new class" of people coming from a different background would inevitably replace their predecessors. Consequently, there would be no predictable end of history.

Culturalism seemed less divisive than its rivals. It helped explain why some societies were more tolerant than others toward difference, why some were risk taking and others risk averse, why some were stable and others unstable, and why most were lagging behind a minority of the rest in economic development and political liberalization. For instance, "thin" culturalism claimed that Africa and India got a late start in the aftermath of independence not because of colonial rule or corrupted elites but because investment was not a value; "thick" culturalism traced the gap between rich and poor nations in these regions to an alleged incompatibility between religion and profit—an argument turned upside down farther East when an incredibly rapid industrial growth was attributed to "Asian values."

In spite of numerous caveats, cultural differences do have deep consequences for institutional design and political action. Beliefs have an impact on individual behavior, and the way peoples solve their internal and external conflicts varies among countries. Memories of conflict resolution frame current decisions. Even false beliefs about reality matter: The misperception of Arabs' capabilities explain the early defeat in 1973 of the overconfident Israelis; the belief that Iraq had weapons of mass destruction contributed to support for an enduring civil war that Western conceptions of cost–benefit analyses did not anticipate; and the French or Americans may well believe in their "exceptionalism."

However, if culture reproduced itself fully from one generation to the next, change would be impossible. Deviants, revolutionaries, and reformists of all sorts—those who do not believe in the core values shared by their fellow citizens—make history, and culturalism is little equipped to examine them. Could neo-culturalism address change better than classical culturalism does?

Toward a New Conception of Culturalism

Each time a particular explanation of political and international life loses attractiveness, a new generation of scholars adapt it to the social and global context of their time. Two contenders succeeded behaviorism: political psychology and rational choice. Functionalism (as in European studies); structuralism, realism, and liberalism (as in international relations); or institutionalism (in policy analysis) were rejuvenated with the addition of the word *neo* to their traditional denomination. Culturalism did not follow suit, and no scholar coined a hyphenated "neo-culturalism" to save it from growing irrelevance. The time has come to argue that rejuvenating culturalism could open new avenues to both "positive" *and* "meaningful" knowledge. This would also preempt the occupation of available academic space by "constructivism" (when it refers to *imagined communities*, *identity building*, etc.) and "cognitivism" (whose *learning processes* look very similar to "socialization" and "acculturation").

Neo-culturalism could be shorn of embarrassing claims that mattered much in the demise of classical culturalism: On top of the list is an insistence on "specificity" that did not bring to its predecessor the benefits historians gained from their own use of "historicity." On the contrary, classical culturalism was deeply affected by critics of its contradictory assumptions that values were either universal or not comparable. Every society is deeply divided and multicultural to some extent: Even the society considered as the most homogeneous in the world, Japan, is composed of several communities of origin and will accommodate more migrants in the future. As evidenced in several comparative surveys, "pride" and "loyalty" are shared less widely across nations than assumed by pollsters. And it is a solid lesson of history that even in the ancient world, "national" cultures never erased local ones despite considerable efforts made by writers and architects to promote the official culture in monuments, poetry, epics, and drama.

Neo-culturalism would also show how diverse were the answers given to the same questions in world history and what these differences actually mean. For instance, public deliberation on policy issues is universally attested in every documented case, but the format of the debate differs from one case to another, and variations matter. The iconography of public deliberation is full of insight about the reasons why a particular configuration was once preferred over other alternatives.

There is also a realm in which neo-culturalism has no rival: tracing individual behavior to interpersonal connections. This would apply first to the records of reciprocity and solidarity obligations and the rights and duties assigned to each actor according to his or her place in a vast system of interactions. It would also show that lineages are at the heart of such networks, be they family links, kinship ties, or relationships with neighbors. After all, the unexpected success of "social capital" stems from its ability to focus on networks of interactions among people.

It may even happen that cultural differences account for variations in social organization, institutional design, and policy making. The most striking example is the preference for endogamy: In Arabic-speaking countries, marrying a first cousin is an ideal, as was the case several millennia ago in the Middle East. Being incompatible with political alliances, it is not conducive to state formation, whereas exogamic marriage prevents military feuds, enlarges property, and makes "foreign" rule acceptable. On the contrary, breaking with blood ties as in the Japanese adoption system through which promising young adults could change family names (*yôshi*) is conducive to peace and prosperity.

Finally, neo-culturalism has much to say about languages as frames of perception of the real world. It casts new light on the meaning of institutions that are usually taken as essentially uncontested. For instance, s*tates* (in the languages derived from Latin or borrowing from it) are presumably "established," whereas *dawla* (in Semitic languages) experience revolutionary cycles. The *koku/goku* conflation of nation and country in Northeastern Asia differs from the Western practice of hyphenating states and nations and is quite telling about the autochthonous versus nonautochthonous conceptions of loyalty. In the same vein, symbolic distinctions, for instance, commemorations and official narratives, were made possible because people forgot the cleavages and hard disputes that gave birth to them, their role being to celebrate a current union carved out of historical divisions.

The Added Value of Culturalism in Political Science

Cultural variables are useful when every other factor fails to give satisfying explanations for a phenomenon. Therefore, its minimum contribution to science lies in its capacity to mobilize residual factors that may have been neglected by other paradigms.

However, neo-culturalism also underscores the role of kinship, heredity, and clienteles in politics. It brings to the fore ethnicity and identity issues without being prejudiced against demands for recognition that may appear to be too far-fetched in normative theory. It certainly says much more about worldviews as conveyed by language and their impact on domestic and international politics.

It is difficult to forecast the future of neo-culturalism in political science. Whereas there is little chance that it will ever benefit from the kind of appeal that boosted classical culturalism in the 1960s and early 1970s, it may engender explanations complementary to more structural and rational theories. Moreover, this paradigm may survive in other guises, such as constructivism (with its insistence on identity and recognition issues) and cognitivism (with its promises to better understand how ideas, inherited attitudes, core principles, and causal beliefs influence perceptions of current issues and solutions to collective problems).

Yves Schemeil
Institut d'Etudes Politiques de Grenoble
Grenoble, France

See also Civic Culture; Discourse Analysis; Ideology; Political Culture; Social Capital; Values

Further Readings

Almond, G. A., & Verba, S. (Eds.). (1989). *The civic culture revisited*. Newbury Park, CA: Sage.
Blondel, J., & Inoguchi, T. (2006). *Political cultures in Asia and Europe: Citizens, states, and societal values*. New York: Routledge.
Diamond, L., Linz, J., & Lipset, S. M. (Eds.). (1994). *Political culture and democracy in developing countries*. Boulder, CO: Lynne Rienner.
Putnam, R. D. (1994). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
Said, E. W. (1979). *Orientalism*. New York: Vintage.

D

Data, Archival

Although any data that are safely stored and retrievable can be described as *archived* or *archival*, here the term applies only to the records of people or organizations that are created for their own use, whether or not they have been collected or manipulated by a researcher. Sometimes the records are numerical. Then, if the data are judged sufficiently accurate, complete, and relevant, they can be compiled and analyzed statistically. The material is more often plain text (or sometimes images) originally created to meet contemporary personal or organizational requirements. This entry discusses challenges in obtaining access to materials, some methodological questions, the impact of digitization of information, and frontiers in archival research.

Access to Materials

Because modern governments and most nongovernmental organizations usually produce a written record of their activity, the variety of archival data and the purposes for which researchers use such data are too numerous to count. Probably the most common objective in academic research on politics has been to analyze high-level government decisions, especially foreign policy decisions. Archived records are often the most important sources of information about these decisions, and sometimes the only one. Even though attempts to perform process tracing on policy decisions can also use interviews or oral histories, analyses of open governmental and nongovernmental sources, statistical analyses of observable actions, and findings from other published research, access to even a small portion of the internal records often provides information not available by other means, and the opening of formerly secret records often leads to significant changes in how decisions are understood.

Researchers who use archival data generally do not regard them as offering an unbiased and complete account of events. Most governments do not operate under laws that establish a presumption of public access to government records. They almost always restrict access to much of their records, especially when they have been recently created, relate to foreign policy, or are politically sensitive for some reason. Although in some countries the law provides that citizens have certain legal rights to government information and can request the release of restricted material, governments can implement these laws in ways that partially nullify their effect by allowing long delays, imposing large fees for search and document production, construing search requests as narrowly as possible, and sometimes simply ignoring their legal obligations. Although determined petitioners with the requisite money, time, or political influence can sometimes overturn these restrictions, they can frustrate some researchers and discourage others. Even when they make available a large portion of their records on a subject, governments often hold back material when its release is deemed to have significant negative political consequences. Governments also destroy a substantial share of the documents they create, and they often redact portions of documents that they release to the public. They might alter their policies on the release of restricted material in response to current political events, in some cases actually reimposing restrictions on material that formerly had been available (this has happened in both the United States and the Russian Federation).

Governments can facilitate or hinder the use of archival material by how they manage their records. Even when documents have been deposited in an archive, they must usually be sorted, catalogued, and listed in a finding aid, checked against mandated restrictions on their full or partial release, and stored in suitable containers. Funding levels and legal or donors' restrictions can greatly affect the rate at which processing can proceed, and fees for photocopying or retrieving material will constrain those without the necessary funding. Archives typically hold large inventories of unprocessed materials; how much value this material has to researchers depends not only on how rapidly it is processed but also on how archivists use any discretion that they might possess to alert researchers to unprocessed material that might be relevant to their research interests. In the United States, the delays in governments' response to many requests for declassification are a significant hindrance to the research for books, articles, theses, and dissertations that rely on such material, as researchers often face deadlines that might be impossible to meet if important material has not been promptly declassified.

Methodological Questions

Very little has been written about the methodology of archival research. A pioneering effort is Marc Trachtenberg's *The Craft of International History*, which discusses search strategies, evaluations of document quality and significance, and how multiple sources can be used to cross-check document contents and researchers' conclusions.

In a country such as the United States, the massive holdings of the National Archives (as of 2003: 2.8 million cubic feet of paper text and 4.7 billion electronic records) require that researchers who intend to investigate this material possess some framework or theory that guides their search. Relying on material used by previous research or presented in government publications as a starting point is common. If one views archival records as documenting the flow of information within a single social network, the choice of starting point might not seem crucial; in principle, all one needs to do is to follow the documentary trail as it takes the researcher first to one decision maker, then to others, until the full network is covered. As a practical matter, several factors often make tracing the information flows quite difficult and comprehensive coverage unlikely. The organization of material is often dependent on the personal practices of specific officials, agencies, or presidential administrations, and learning where to search is not always quick or easy. Especially sensitive information might never be written down. Even when it is, it might be stated in a blind memo or worded such that is very difficult to understand without very good knowledge of the historical context. It is sometimes difficult to know whether an official has actually read a document—as opposed to him merely being on the distribution list. Finally, the time and resources available generally are insufficient to examine all material, so some sampling procedure is usually followed, even if only implicitly.

The interaction of the sampling method with the decision to begin with the material used in previous research involves some risk. For example, in studying diplomatic history, a presupposition and a consequence of archival research that follows such a strategy is that the information collected will say relatively little about officials' interaction with private persons. Government publications of archival material ordinarily include very few items that were not created by government officials. Although archives often contain correspondence files that contain letters from private persons and less often records of office appointments, telephone calls, or other interactions with nongovernmental contacts, it might require a fair amount of contextual knowledge to appreciate which contacts are significant and which are not. The impression that such interactions have little impact is further strengthened by a widely followed practice of writers of governmental documents, who typically justify and explain policies in terms of plausible "public interest" claims. This strategy is sensible because such justifications by definition will have a wide appeal, and because considerations of narrow political advantage are often viewed as illegitimate motivations for government policy. While scholarship that emphasizes the role of interpretation in the reading and writing of texts has become one of the main intellectual currents within the social sciences over the past quarter century, its concern with "authorial positioning" and interpretation as a choice rather than a reflex seems to have been focused more often on those who read and write about the documents than on those who create them.

Source evaluation and interpretation is also a challenge that can typically be met only with a good understanding of the context within which material was created. Before investigating an archive, one usually has acquired some background knowledge from previous research. At a minimum, this provides a list of the most visible participants in the formal decision-making process, their most visible concerns, and some conception of the flow of events. Information about informal mechanisms and more sensitive (and thus more private) concerns and information is valuable in opening a path to understanding the less obvious factors bearing on a decision. As Trachtenberg explains, prior understandings held at the outset of archival research are generally reshaped by the encounter with new material. This creates a dynamic process in which interpretation of new material triggers changes in the interpretation of material previously collected, which then alter search patterns or the interpretation of additional new material. As the amount of potentially relevant material is often vast and cannot be interpreted by any one individual, this process is a social rather than merely a personal activity.

Impact of Digitization of Information

The widespread computerization of records and the shift to electronic communications such as e-mail and texting has had complex effects on archival research. Making document images or files available on the Internet makes them much more accessible, but only a very small proportion of government and nongovernment records are available this way. Having finding aids and catalogues available online is a significant advance, making it possible to learn in detail about archival collections without needing to visit. If documents were originally created in electronic form, archival storage of the documents is facilitated, provided durable and stable storage media are used. If text is machine readable, then it is much easier to search it systematically, analyze it with text analysis software, and excerpt it for teaching or publication purposes. However, whether governments will decide that e-mails, texts, and other electronic communications are as "real" as paper documents, and devote substantial effort to preserving them, is difficult to predict.

Archival Research

Trachtenberg is the first scholar to present systematic "declassification analysis," which inspects streams of declassified material on a topic over an extended period to infer the basis for declassification decisions from the characteristics of what was released and withheld at various times. This practice also illuminates what material a government believed to be especially sensitive, though that understanding is, of course, heavily context dependent.

Text analysis software is now beginning to be used in archival research. Even ordinary word processors can facilitate analysis by searching text databases for strings of text. Specialized software provides for more complex search strategies (such as Boolean searches), attaching annotative tags to text, and relatively easy and rapid recombining of document excerpts.

Government information policy is a topic that has received little scholarly attention in political science. In view of its pervasive effect on how citizens and researchers understand government, it is a topic that should receive significantly more attention.

Timothy J. McKeown
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina, United States

See also Data, Textual; Discourse Analysis; Process Tracing; Quantitative Methods, Basic Assumptions

Further Readings

Burton, A. (Ed.). (2005). *Archive stories: Facts, fictions, and the writing of history*. Durham, NC: Duke University Press.

Cox, R. J., & Wallace, D. A. (Eds.). (2002). *Archives and the public good: Accountability and records in modern society*. Westport, CT: Quorum Books.

Federation of American Scientists, Project on Government Secrecy: http://www.fas.org/sgp

Gidiere, P. S., III. (2006). *The federal information manual: How the government collects, manages, and discloses information under FOIA and other statutes*. Chicago: American Bar Association.

Trachtenberg, M. (2006). *The craft of international history: A guide to method*. Princeton, NJ: Princeton University Press.

DATA, MISSING

Missing data problems arise frequently in social science applications. Household survey respondents may fail to provide responses to questions for reasons of refusal, fatigue, enumerator error or perhaps a split sample design is used, in which case the missingness is an intentional strategy to conserve resources. Personal income is often missing in surveys due to respondents' refusal to reveal what they earn. Cross-national data on macroeconomic conditions may fail to include entries for some countries in some years due to crises, lack of capacity to generate data regularly, or *ex post* realization that past reporting methods were flawed. Measures of national income inequality are often missing for some countries in some years in cross-national data because their calculation requires an income survey; such surveys are too expensive for some countries to implement every year. In the household survey example, the analyst must decide whether and how to use data from respondents who refused to state their income. In the cross-national example, the same issue arises for country-years for which the inequality measure is missing. These are instances of what analysts may label as "item missingness." This is the focus of this entry.

The "items" are the individual survey questions or bits of information. The idea is that for each household or country, analysts have information on some items but not for others. This is to be distinguished from "unit missingness," where "units" refer to units of observation—the households or country-years in the examples above. That is, unit nonresponse refers to the situation in which data collection failed to accumulate *any* data for some units, for example, some households were unreachable or refused to participate altogether, or some country-years contained no usable data on any variables. Fixes for unit nonresponse in surveys include *ex post* weighting (via post-stratification) to externally available population information, such as a census. For cross-national data, the analyst typically has no way to adjust for unit missingness, and so he or she is required to state that the results hold only for the subpopulation for which data are available.

Listwise Deletion

By far the most common approach to missing data in political science has been to simply drop incomplete cases, something known as "listwise deletion" in the methodological literature. This is tantamount to taking cases of item missingness to be instances of unit missingness. In special circumstances, ignoring cases exhibiting item or unit missingness is a perfectly reasonable thing to do. For a unit i, consider an outcome variable—call it y_i—and a set of predictor variables recorded for the unit—call them x_i, collectively. Suppose that the distribution of y_i depends on x_i in a manner characterized by a probability density function $f(.)$. In that case, the analyst can write the probability density of y_i given x_i as $f(y_i|x_i)$. An example would be a linear structural model of the form $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$, with $\varepsilon_i \sim N(0, \sigma^2)$. In this case, $x_i = (x_{i1}, x_{i2})$, and

$$f(y_i|x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2})^2}{2\sigma^2} \right).$$

The goal of data analysis in such a circumstance is to estimate the parameters that characterize $f(y_i|x_i)$—here, the β coefficients and the variance, σ^2, of the error term, ε_i. Listwise deletion allows the analyst to do so without bias as long as two conditions hold: (1) the analyst has correctly specified the functional form of $f(y_i|x_i)$, and (2) the probability that an item for i is missing systematically depends only on the elements of x_i that are always observed for all units.

This is an important result, but one should understand that Conditions 1 and 2 are never exactly true in social science data analysis. In social science, analysts work with approximations (e.g., regressions with linear or maybe quadratic terms) of the complex and unknown functional forms that govern the relationships in their data. This violates Condition 1. The result for unbiasedness does not hold for estimating parameters in an approximation model. Missingness may depend on the value of y_i or the value of the missing item itself (if other than y_i), violating Condition 2. Therefore, using listwise deletion requires a leap of faith: Analysts assume that the inevitable bias is small and thus of little importance. This may or may not be true.

Even when bias is not a major concern, listwise deletion can be highly inefficient. It can cause analysts to discard lots of information that could improve the certainty of their inferences about parameters of interest. To see this, suppose that an analyst has a data set with five variables that are needed for the analysis. Suppose each variable has a 5% chance of being missing. The analyst expects to have recorded information 95% of the time for each variable, but listwise deletion implies that about $(1 - .95^5) \times 100\% = 23\%$ of the observations will be discarded. This is a sizeable amount of data to discard, making inferences considerably less precise. The analyst would be throwing away many cases with only one missing item, which is wasteful. There are better ways to proceed.

Multiple Imputation

This entry does not discuss what is known as "selection bias," where missingness occurs only on an outcome variable that cannot be observed or where missingness is a result of actions of actors that should be modeled. Such missingness occurs, for example, when analysts do not observe the votes of those who abstain or the outcomes of wars that states decide not to engage in. The remainder of this entry is concerned with cases of missing data for items that are either designated as "predictor" variables in a regression analysis, can sometimes serve as predictors and sometimes as outcomes, or will be used for descriptive analysis (e.g., to estimate a population mean). The presentation is based on Rod Little and Donald Rubin's seminal book *Statistical Analysis With Missing Data*, first published in 1987.

Imputation would seem to offer a solution. Imputation refers to "filling in" (i.e., "imputing") missing items with reasonable estimates of what those items should be. Simple imputation methods include selecting item values from similar units (what is called "hot deck" imputation), imputing the mean or median value, or using a regression to predict a conditional mean. Census bureaus and other statistical agencies employed such simple methods until recently. As Little and Rubin have helped clarify, such simple imputation methods can create problems. First, they distort covariance relationship between variables in a data set, introducing bias into subsequent analysis. Second, imputing a single item and then treating it as the "true" item value for that unit ignores one's uncertainty about whether or not the imputed item value is correct.

For these reasons, methodologists have developed "multiple imputation" as a statistically sound way to address item missingness in a principled manner. This method takes care to preserve the overall covariance structure of the data set by, essentially, adding the right kind of noise to the imputations. Uncertainty about the imputed items is propagated by generating multiple data sets in which missing items are imputed. The substantive statistical analysis is then conducted on each of these data sets, and the results are averaged. Standard error estimates incorporate both within-sample variability and the variability across the imputed data sets. The method and accompanying theory were developed in Rubin's 1987 work, *Multiple Imputation for Nonresponse in Surveys*.

Multiple imputation requires that either missingness is completely random or variables exhibiting missingness can be modeled well using other variables in the data set. This is an instance of what statisticians refer to as "ignorability." This is not true when the missingness pattern in a variable, y, depends on the values of y itself in such a manner that other variables in the data set are of limited use in modeling this pattern. The latter case is called "nonignorable" missingness and can be corrected only by making non-innocuous assumptions about how the missing data behave.

Many multiple imputation algorithms have been incorporated into commonly used statistical software packages. The two most common types of algorithms are those that use (1) a multivariate model of the entire data set or (2) a series of regressions. An example of the first type is the Amelia algorithm, first introduced by Gary King, James Honaker, Anne Joseph, and Kenneth Scheve. Their article, "Analyzing Incomplete Political Science Data," may be credited with doing the most to make political scientists think about using multiple imputation rather than listwise deletion. Amelia models the entire data set with a multivariate normal model. It uses a novel iterative approach to estimate all of the mean and covariance parameters for the model and fills in the missing data with random draws from the model. Regression-based algorithms are based on the following idea: For a variable, y, exhibiting missingness, a regression model using the other variables in the data set is specified. Then, the parameters of the regression model are estimated, including, importantly, the variance of the random errors and terms characterizing uncertainty about the parameters of the model. Then, missing values on y are filled in with random draws based on the fitted model and incorporating all sources of random variability in the model and the parameter estimates. Because the regression model will generally include predictors that themselves exhibit missingness, an iterative sampling algorithm is used to fit the models. Regression-based algorithms may alternatively use "predictive-mean matching," which allows the analyst to avoid having to model the data directly. Predictive-mean matching uses linear approximations to match units that exhibit missingness to "donor" units that do not exhibit missingness. Missing values are filled in with the values from the donors.

Other alternatives exist in principle for resolving problems of item missingness. These include maximum likelihood or Bayesian methods that skip the imputation process and directly adjust the substantive analysis by taking missingness patterns into account. Such approaches are rarely used in practice, however, because they require significant application-specific statistical programming. Such methods have not been shown to perform any better than multiple imputation, although, as mentioned, the number of attempts are few.

Cyrus Samii
Columbia University
New York City, New York, United States

See also Regression; Selection Bias

Further Readings

King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review, 95*(1), 49–69.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Data, Spatial

Data of interest to a political scientist often refer to attributes of units—for example, individuals or nation-states—that have a spatial location. Although analysts typically disregard geographical position and spatial configuration and treat their data as collections of independent units, there are many reasons why it may be essential to consider the spatial ordering of the data. In particular, attention to the spatial dimension can help generate new insights and may be essential for drawing correct inferences about the influence of other features in statistical analyses, even when spatial patterns are not the researcher's primary concern. In the following, the impact of spatial relations and their use in regression analyses are discussed.

Spatial Relations

One way to appreciate the importance of spatial patterns in data is to consider the analogy to time-series data. Researchers are increasingly sensitive to how individual data points collected over time often will not be independent of one another, as many social and political processes display clear trends and persistence over time. To use a concrete example, consider how people's evaluations of an incumbent government typically change slowly over time. Data on approval taken at two points that are close in time are thus likely to display similar values or serial correlation. It would be misleading to consider the individual observations a random sample, and analysts will need to take into account their temporal ordering if they want to examine how other features such as economic performance may influence changes in approval. If economic performance is believed to influence approval and such data also display a trend over time, then analysts risk making spurious inferences if they do not take into account the possibility of serial correlation.

Social science researchers have been much less attuned to the possibility that data collected for units at the same point in time may not be independent of one another across space. Francis Galton pointed out in an early influential comment on an article comparing marital institutions across societies that the observed outcomes could result from diffusion between societies rather than independent processes operating in isolation in each unit. Stated differently, common marital institutions may not reflect a functional relationship to certain shared social characteristics, as implied by the author of the article Galton commented on, but could stem from a diffusion process in which societies may be likely to adopt the institutions found in other societies that they interact with. Distinguishing between the results of independent outcomes within each unit and diffusion may be difficult, hence the term *Galton's problem*. For example, societies that interact more with one another may be similar in other respects, such as in economic or other social structures. Analyzing the observations as independent units and looking for functional relationships between social characteristics and marital institutions would run the risk of making unwarranted inferences about causal relationships from social characteristics to marital institutions, unless the possibility of diffusion or spatial dependence among units is taken into account.

There are many reasons to suspect that political and social phenomena often reflect diffusion or spatial dependence, in the sense that observations that are "near" or "connected" in some manner tend to have similar values. For example, individual attitudes may not just reflect an individual's own characteristics but could also be influenced by the people that this person interacts with. As such, two individuals who are connected by common ties may have similar attitudes, even if they at a first glance would seem very different on other social characteristics. Likewise, policies and institutions may be influenced by the policies and institutions adopted by other states. For example, the prospects for democratic reform in autocracies may be influenced not just by domestic events and characteristics but also by international ties and events in other states. Likewise, certain policies such as smoking bans have evolved from being perceived as an extreme form of legislation to become a mainstream policy adopted by many states with very different characteristics. Beth Simmons, Frank Dobbin, and Geoffrey Garrett provide a typology of different diffusion mechanisms for policies and institutions, distinguishing between *coercion*, *competition*, *learning*, and *emulation*. Attention to spatial dependence can be very helpful in learning new insights or developing new hypotheses about the phenomenon of interest. However, it is unlikely that a single theory of diffusion or spatial dependence will be generally applicable, and such theorizing is probably best done on a case-specific basis.

More generally, there is spatial clustering or dependence if the values of a response variable for two units that are linked or connected tend to be similar, so that positive/negative deviations from the sample average for one unit are likely to go together with similar deviations from the average for the other unit. The map of autocracies, anocracies, or democracies in 2004, as classified by the Polity data (Figure 1), provides an example of spatial dependence, in the sense that democracies are much more likely to have democratic neighbors and autocracies are much more likely to be surrounded by other autocracies.

The consequences of spatial dependence for statistical analysis are similar to the consequences of serial correlation over time. Although spatial correlations in principle could be both positive or negative, nontrivial spatial correlations in the social sciences tend to be positive, and this section focuses on this case (although one could imagine negative correlation arising, e.g., through free-riding behavior, where higher spending by

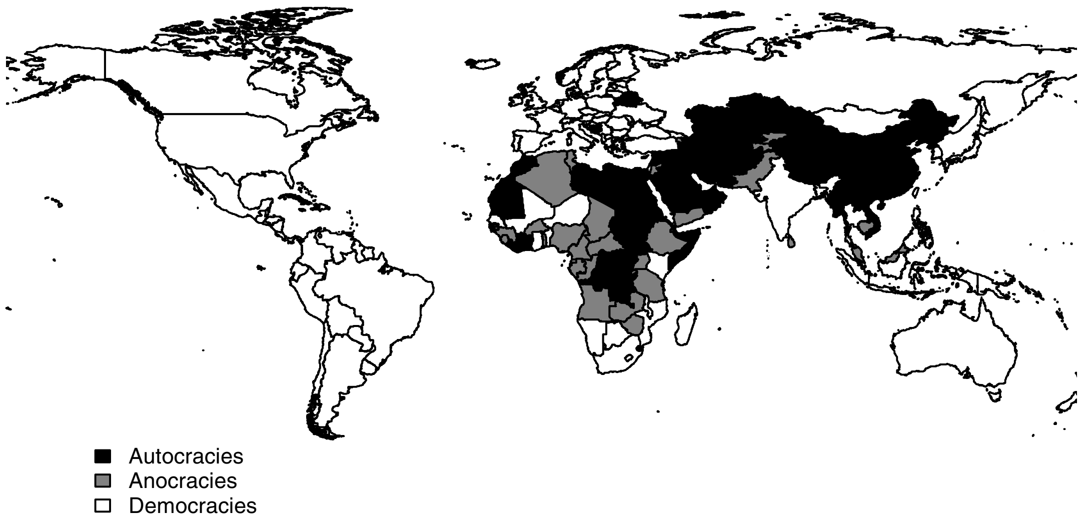


Figure 1 Political Regimes in 2004 According to Polity Data

one unit makes other connected units likely to spend less). In a linear regression setting, the coefficient estimates are no longer efficient estimates of the true parameters if one ignores spatial dependence and could be substantially off the correct values. Moreover, the standard errors assuming independence are incorrect and generally too small in the presence of positive spatial correlation. A helpful, intuitive way to understand why standard errors are likely to be too small in the presence of spatial correlation is to consider its impact on effective sample size. When observations are serially correlated, either in time or in space, so that close observations are similar to one another, the sample does not contain as much independent information as would be suggested by the apparent sample size N. The normal standard error estimate ignores the correlation between nearby observations and will thus understate the true variance.

To detect whether data display spatial clustering first requires the specification of what is meant by "space." In general, the structure of spatial dependence must be prespecified by the analyst, and it is only possible to test for a specific form of spatial dependence. Whereas time has an inherent ordering and data are usually collected at intervals that suggest likely temporal dependence structures, space is multidimensional and there may be many possible forms of dependence between observations. Since much interaction tends to be geographically determined so that units that are close to one another tend to interact more, geographical distance is often considered. However, one can imagine forms of dependence other than geographical distance being important. For example, observed indicators of interactions such as trade could be used to specify likely patterns of dependence, or one could consider links between units that are similar on some other criterion, such as a shared language.

Regression Analysis

Given some criteria for potential spatial dependence, the analyst can represent the potential structure of dependence as a graph, specifying how units are connected to one another, or as a connectivity matrix $\mathbf{W}$, where the value of an entry w_{ij} indicates whether (and to what extent) one unit i is linked or connected to another unit j. Connectivity matrices can represent a very wide range of possible dependence structures. The connectivity matrix may be based on a single threshold, where two units i and j are considered connected, and their corresponding entries w_{ij} acquire nonzero values if they meet this threshold. Alternatively, the connections could be weighted—in the sense that units are "more" connected the closer two units are on some criterion—for example, with values based on inverse decay of the distance between units. A connectivity matrix could be symmetric, to reflect a reciprocal relationship, where a link from *i* to *j* goes together with a link from *j* to *i*; or it could be directed in an asymmetric manner, so that the fact that *i* influences *j* does not necessarily imply that there is a reverse link from *j* to *i*.

A variable y displays spatial correlation to the extent that the value of y for one individual unit i is similar to the values of y for proximate or connected observations j. Since a single unit usually has many connected units, one must take into account the degree of similarity across the connected observations to assess the overall spatial correlation. For many purposes, it is useful to row-normalize the connectivity matrix so that all the connectivities for a single unit sum to 1. This is often referred to as a standardized spatial weights matrix **W** in the spatial econometrics literature. This allows one to generate a "spatial lag" of a variable y by postmultiplying it with the weights matrix **W**—that is, **W**y. The individual elements of the resulting vector for an observation i can be interpreted as the average of y among neighbors or the other observations j that are connected to i following the specification of the connectivity matrix **W**. A variable displays positive spatial correlation if the values of y for one unit i, y_i, are similar to the values for its connected observations given by row i of the connectivity matrix—that is, $w_i y$.

One useful summary measure for spatial correlation of a variable y is Moran's I correlation coefficient. This is based on comparing the deviations from the mean of y for a unit i with the deviations from those of its neighbors—that is,

$$I = \frac{N\sum_i\sum_{j,j\neq i} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_i\sum_{j,j\neq i} w_{ij}\right)(y_i - \bar{y})^2}.$$

A standard error can be approximated for Moran's I, based on sampling assumptions, which in turn allows testing whether the spatial clustering is significant or deviates from what would be expected if there were no spatial pattern. Moran's I can be given a useful graphical interpretation. Figure 2 shows a plot of a variable (in this case the Polity democracy scale in 2004) on the horizontal axis, standardized to have a mean of 0 and a standard deviation of 1, against its spatial lag on the vertical axis (here based on taking countries

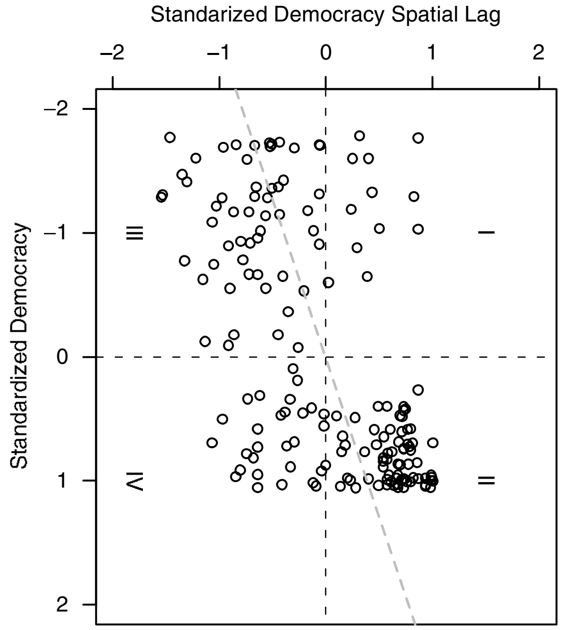


Figure 2 A Graphical Illustration of Moran's I, Based on Polity Data for 2004

within 500 kilometers of the outer boundaries of a state as connected to one another). Drawing vertical and horizontal lines through the origin (0, 0) divides the figure into four quadrants, where the values of the individual observations variable are either below or above the mean on the input variable and its spatial lag. If there is positive spatial correlation, one would expect to see positive deviations from the mean on the input variable for one unit to systematically go together with positive deviations from the mean for its neighbors or the spatial lag. There should be relatively few observations far from the dashed lines of the averages in the upper left (I) and lower right (IV) quadrants, where the sign of the two deviations diverge, and the bulk of the individual observations should be in Quadrants II and III, where positive and negative deviations go together for both variables. Although some cases fall inside the boundaries of Quadrants I and IV in Figure 2, the bulk of the observations are clearly in Quadrants II and III, reflecting how most countries tend to have levels of democracy similar to the weighted average of the levels found among their neighbors. Moran's I is the slope of the regression of the spatial lag on the standardized input variable. This is indicated in Figure 2 as a gray dashed line, and one can see that the slope of the regression line clearly is positive. In contrast, if there is no spatial correlation, one would not see any tendency for the deviations to have similar signs, and the slope of the regression line would be close to 0. Moran's I statistic can be used to detect both spatial clustering in an input variable and spatial correlation in the residuals from a regression, and this is implemented in common software options (see Figure 2).

When theory or preliminary analysis suggests spatial dependence, researchers may wish to try to integrate the spatial component in a regression model. The appropriate model specification depends on specific theoretical arguments or the likely causes of spatial dependence as well as on whether the dependent variable is continuous or categorical. Model estimators for spatial regression models with continuous dependent variables are relatively well established, and this section focuses only on this case. Models where the dependent variable is categorical, such as binary responses or counts, are considerably more complex, but various estimation approaches relying on simulation or data augmentation have been proposed.

As in the case of time series, one might approach spatial dependence either as a nuisance to estimation that must be taken into account when considering other phenomena of interest or as a phenomenon of interest in its own right. In the former case, researchers may find it helpful to consider a regression model where the errors display a spatially correlated structure. This can be written as $y = \mathbf{X}\beta + \lambda\mathbf{W}\zeta + \varepsilon$, where $\mathbf{X}\beta$ denotes the systematic component of the model, $\mathbf{W}\zeta$ indicates a spatially correlated error structure as specified by the connectivity matrix $\mathbf{W}$, and ε is a residual random error. The parameter λ reflects the extent to which errors are spatially clustered. If $\lambda = 0$, then the model reduces to a standard linear regression, where the observations are independent of one another. If $\lambda \neq 0$, there is a spatially clustered pattern, where a linear regression treating the observations as independent of one another may give misleading results. This approach may be appropriate if a researcher believes that there is likely to be spatial dependence arising from some spatially clustered omitted variable. If the model is correct, the consequences of spatial dependence for inference can be addressed by letting the errors of the model be spatially clustered, so that the sign and size of errors for individual observation covary spatially. This can be seen as analogous to time-series–correlated error models, in the sense that observations are interdependent only due to omitted variables that are spatially correlated. Once removed from the systematic component to the error, space no longer matters.

One interesting alternative model of spatial dependence from a substantive perspective is the so-called spatially lagged dependent variable, which can be written as $y = \rho\mathbf{W}y + \mathbf{X}\beta + \varepsilon$. The dependent variable for connected units itself is here specified as having a direct impact on the value of the dependent variable for unit i and appears on the right-hand side of the model. The parameter ρ indicates the extent to which the value of the dependent variable y for a unit i is influenced by the values of y of its neighbors j. This can be seen as a spatial analogy to a time-series model with a temporally lagged dependent variable. The spatially lagged dependent variable model is straightforward to estimate as a standard linear regression if the spatial impact operates with a time lag, so that the values of y for the connected units affecting i can be taken as predetermined at time t. However, if the spatial impact is considered to be contemporaneous, or within the same period, there is a problem of simultaneity, since y at time t appears on both the left- and the right-hand side of the equation and the individual values are jointly determined. The model must be estimated via either instrumental variables or a maximum likelihood estimator, which is now implemented in many common software options. Standard significance tests can be used to test whether the spatial lag has a significant impact on the response variable, and the model with the spatially lagged dependent variable can be compared with a model assuming independent observations using standard comparisons and tests of nested models.

The interpretation of a model with a spatially lagged dependent variable differs notably from that of a model where the units are considered to be independent. In the latter, the coefficient estimates β for a right-hand-side independent variable x can be interpreted as the expected change on the dependent variable y following a one unit change in x. However, this interpretation no longer holds in a model with a spatially lagged dependent variable, since the model implies spatial feedback. The net effect of a change in x for i will now depend on how the effect of the change of x on y for i influences the dependent variable for other units j connected to i, which in turn will feed back on to the value of y for i. Assessing the net effect of a change in x for i must take into account not just the immediate effect β that increasing x will have on i but also the long-term effect of the change on y for unit i on the value of y for connected observations and the implications that these will have for other units connected to these until the system settles on some new equilibrium. If the spatial feedback implications are large, the net effect of a change in x could be considerably larger than the immediate impact given by β. Through rearranging, it is possible to show that the expected effect of a change in x is given by $(\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\beta$, where $\mathbf{I}$ is an identity matrix. This tells one that the specific impact depends not just on the first-order connectivities stipulated in the connectivity matrix $\mathbf{W}$ but also on the higher-order connectivities implied in the spatial multiplier $(\mathbf{I} - \rho\mathbf{W})^{-1}$. Since units i can have different degrees of connectives to other observations and different degrees of influence on other observations, the effects of changes in x will be unit specific and may vary considerably across units.

Applications of spatial dependence in the social sciences have until recently been fairly uncommon, but the past decade in particular has seen a wealth of new applications appearing in print. These applications have been facilitated by a renewed interest in interaction and networks in social science theory as well as the increased availability of software to conduct spatial analyses, including geographic information systems resources and routines for general statistical packages such as MATLAB, R, or Stata.

Kristian Skrede Gleditsch
University of Essex
Colchester, United Kingdom

See also Aggregate Data Analysis; Comparative Methods; Maximum Likelihood; Network Analysis; Regression; Sampling, Random and Nonrandom; Statistics: Overview; Variables, Instrumental

Further Readings

Beck, N. L., Gleditsch, K. S., & Beardsley, K. C. (2006). Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly, 50*(1), 27–44.

Bivand, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2008). *Applied spatial data analysis with R*. Berlin, Germany: Springer.

Cressie, N. (1991). *Statistics for spatial data*. New York: Wiley.

Simmons, B. A., Dobbin, F., & Garrett, G. (Eds.). (2008). *The global diffusion of markets and democracy*. Cambridge, UK: Cambridge University Press.

Ward, M. D., & Gleditsch, K. S. (2008). *Spatial regression models*. Thousand Oaks, CA: Sage.

Data, Textual

Textual data refer to systematically collected material consisting of written, printed, or electronically published words, typically either purposefully written or transcribed from speech. Text collected for use as data typically reflects a conscious research purpose, motivated by a design aimed at yielding insight on some feature of the social or political world. This entry outlines the purpose, issues, and challenges involved in selecting, preparing, and analyzing textual data.

The process of textual data analysis follows several main steps. First, a text and its author are identified that will directly inform the research question at hand, for instance, legislative speeches if one is interested in lawmakers' policy agenda. Second, this text must be processed, typically involving numerous decisions about how to convert, store, edit, and combine texts. Next, the processed text is analyzed, a step that may take any of a wide variety of forms. Humans may read the text to classify it, possibly after dividing texts into smaller units; computerized tools may analyze the text to count frequencies of preidentified words or phrases; statistical methods may be applied to specific words, word patterns, or sentences to construct scales, clusters, or to identify other patterns. Finally, the results of the analysis are summarized, interpreted, and reported to inform the research question that motivated the analysis of the textual data.

The steps in this process are not dissimilar to those found in any empirical research design, but the systematic use of textual data may present special challenges in the stages related to identifying the data to be collected and preparing the data for analysis. In the modern era of near-ubiquitous electronic textual content, this problem stems not from a lack of availability of textual data or tools for working with them but, rather, from the opposite: The staggering quantity and variety of textual data create real challenges in selecting which texts to analyze. Furthermore, the technical challenges of working with different formats for recording electronic texts, as well as converting these into formats or "data sets" that suit the purpose of analysis, are often far from trivial. Despite such challenges, however, textual data remain one of the most promising and one of the least explored sources of systematic information about the political and social world.

The Appeal of Textual Data

Textual data in the political and social sciences have several principal advantages. The first can be found in the nature of text and the information it may contain about the authors or speakers generating that text. If the task at hand is to gauge overt messages or signals conveyed through text, of course, then no better source of information exists than the texts themselves. For instance, in 1951, in a well-known study of articles by other Politburo members about Stalin on the occasion of his 70th birthday, Nathan Leites, Else Bernaut, and Raymond Garthoff were able to discern differences in groups with regard to communist ideology. In this instance, the messages signaled not only an underlying orientation but also a degree of political maneuvering with regard to a leadership struggle following the foreseeable event of Stalin's death. The messages are themselves significant, and these could only be gleaned from the public articles authored by each Politburo member, written in the full knowledge that they would be reprinted in the party and general Soviet press. Similar content is present in the overt messages from advertising—although such analyses frequently focus on nontextual content as well—as well as in political speeches and political advertising. The principal advantage of textual data is that what political actors write or say may be a significant political signal or act, a form of purposive political act best judged using the textual data as the best manifestation of the act itself.

A second notable advantage of textual data for political analysis lies in the possibly unique capacity of text to inform us about the qualities of political and social actors that are unobservable through direct means. The most frequently analyzed quantities of this type in political science are political "ideal points" referring to an actor's preference on left–right policy scales, such as a relative preference for socially and morally liberal policies versus conservative ones. Other preferences could include being relatively for or against a specific policy, such as the repeal of the Corn Laws in Britain in 1846; being for or against further European integration during the debate over the Laeken Convention; or being for or against a no confidence motion. The preferences are inner states of political actors, whether these are legislators, parties, delegates, or candidates, and hence cannot be directly observed. The most common alternative approach not based on texts is to infer preferences from observed voting patterns. Yet it has been widely shown that voting is a highly strategic political act, subject in particular to strict party discipline in most contexts, and therefore reflects preferences strategically rather than sincerely. Texts and speeches, by contrast, are less subject to political monitoring and punishment than voting, since their content in democratic systems tends to be relatively unregulated as well as having fewer direct consequences for policy outcomes. In sum, textual data frequently contain important information about orientations and beliefs for which nonverbal forms of behavior may serve as poor indicators. Short of reading the minds of political and social actors to find out what they are really thinking—something that neither technology nor ethical codes of research make possible—the next best step is to read what they are writing and saying.

A final advantage of text is its ready availability for analysis. Talk is cheap not just in the political sense but also in the sense of being readily and, in most cases, freely available in electronic formats easily amenable to systematic analysis. Political actors leave a trail of recorded text behind them following almost all of their official verbal activities: speeches, debates, websites, party platforms, press releases, all easily accessible usually for minimal cost and effort. Recorded text never ages, and never forgets, unlike, for instance, participants in a panel survey or an expert interview, and can therefore be used to analyze historical episodes as well as contemporary ones. Text never gets tired or worn out, no matter how many times it is analyzed. While these are properties shared by other behavioral data, such as voting records, they are decidedly not features involving human recall, participant or expert interviews, or surveys—other (indirect) methods frequently used to assess the unobservable, inner orientations of political actors. For practical reasons then, textual data offer immense advantages of availability and durability, provided that valid and reliable means can be deployed to unlock their secrets.

Substantive Decisions in the Analysis of Textual Data

Any research using textual data begins with the researcher identifying the corpus of texts relevant to the research question of interest. From this corpus, texts will be sampled or selected for analysis. Texts are generally distinguished from one another by attributes relating to the author or speaker of the text, perhaps also separated by time, topic, or act. A week's worth of lead daily opinion pieces from *The New York Times*, for instance, could form a set of seven distinct texts. A set of debates during a no confidence motion in the British House of Commons could form a set of numerous distinct texts differentiated by speaker. The set of election manifestos from 1948 to 1997 of the Irish Fianna Fáil Party could form a set of texts differentiated by election date. In each case, distinguishing external attributes attached to each text allows researchers to partition the verbal data they contain, thereby distinguishing one text from another. Knowing that a single text represents an identifiable, purposeful verbal act gives it meaning as well as a set of variable attributes that can be used to compare that text with another. This does not necessarily mean, however, that in analyzing the text, these distinguishing attributes will form the basis for analyzing these data. Here it is useful to distinguish the *sampling units* of textual data from the *units of analysis* for analyzing these data. For example, while texts are generally stored as "Bush's 2007 State of the Union Address" and "Bush's 2008 State of the Union Address," these textual data will usually be analyzed by breaking them down into constituent textual units. The most common constituent units of analysis for textual data analysis are words, consecutive word sequences, and sentences. Determining which are most suitable will depend on the research question, the nature of the analytical technique, and sometimes the length and nature of the texts themselves. Knowing how the sampling procedure for the textual data selection relates to the sampling units and the units of analysis may have implications for subsequent inference, given that the units of analysis are not randomly sampled textual data, irrespective of the sampling units.

In many political science applications using textual data, the "sample" of texts may, in fact, be all possible texts generated by the political world for that application. In tracking the words spoken on abortion per day in the U.S. Congress, for instance, a study might examine every spoken utterance in the Senate from 1997 to 2004. Yet even in such situations where a researcher may not confront overt sampling decisions, such as how many newspaper articles to select from which set of days, from a larger population that is too large to cover in its entirety, it is still important to be aware of selection issues that shape what sort of text becomes a recorded feature of the social system. Such "social bookkeeping" has long been noted by historians seeking texts to gain leverage on events long past, but it may also feature in many forms of observed political text, especially spoken text in structured settings such as legislatures. The key is to be aware of the mechanisms governing the generation of text, with an aim to making sure that the observable text provides a valid representation of the phenomenon that it will be used to investigate.

Practical Issues in Dealing With Textual Data

Text comes in many shapes, sizes, and forms, but for text to be useful as data, it is almost always converted to electronic form prior to analysis. Electronic text has the advantage of being easily stored and easily manipulated. Fortunately for modern users of textual data, new texts are almost invariably recorded in electronic form that can be easily converted for analysis. Websites, electronic documents, text stored on optical media, electronic mail, word processing files, news feeds—all provide text that can be easily captured and manipulated for analysis. The exception tends to be text generated from interactive, spoken activities such as interviews. In such cases, interviewers may need to transcribe textual data or use speech recognition software (combined with a thorough human check on the software's transcription) to convert spoken text into written, electronic text.

One potentially tricky practical issue when dealing with textual data concerns the manner in which text is electronically encoded by computers. Broadly speaking, "text" means the representation of language as a set of recorded characters, following language-specific rules for syntax, grammar, and style to be meaningful. Pre-electronic textual data were typically series of characters drawn or printed on paper. For text to be stored electronically, however, each character must be encoded in a way that corresponds to a digital format used by computers. Morse code, for instance, encodes characters as a series of dots and dashes. Unfortunately—but hardly surprisingly—different computers encode character sets in different ways, and this can create challenges when preparing texts for processing together when these texts use different encodings. For instance, two instances of *démocratie* may be considered as two different words by software designed to compile a word frequency table, when the term occurs in two different documents stored with different text encodings. Encoding issues are further complicated when texts must be converted from other formats, the most common being Microsoft Word, HTML, or Adobe's PDF format. The same Microsoft Word document converted to text twice, once in UTF-8 (an implementation of Unicode) and once in Microsoft Windows' 1,252-character set for Western European languages, will not be equivalent if the document includes words with diacritic marks (such as *démocratie* or *Länder*) or even specialized formatting or punctuation marks—such as "smart quotes" or long-dashed elements—contained in nearly all formatted text, including this sentence.

Many users of textual data find it useful to process raw text prior to analysis, with an aim to identifying more appropriate units of analysis than is offered by the raw text. The most common types of preprocessing are reduction of words to their word stems (or "lemmas," which strips words to their stems while also considering context) or elimination of words through the use of "stop lists" or based on their relative infrequency. The first form of textual data preprocessing treats words as equivalent that differ only in their inflected forms, so that, for example, the different words *taxes*, *tax*, *taxation*, *taxing*, *taxed*, and *taxable* are all converted to the word stem "tax." The second common textual preprocessing practice is to remove words that are considered unlikely to contribute useful information for analysis. These words, commonly called "stop" words, are usually function words such as conjunctions, prepositions, and articles that occur in the greatest frequency in natural language texts but add little specific political meaning to the text that would be deemed useful to analyze from textual data. These stop words may be identified in advance in the form of a "stop word list"—on a language-specific basis of course—or may be identified by their relative frequencies across and within sampling units (documents). The problem with excluding words from an a priori list, however, is that there may be no universal, cross-applicable list of words known to contribute nothing useful to all textual data uses. For instance, the pronoun "her," it has been demonstrated, has a decidedly partisan orientation in debates on abortion in the U.S. Senate. For these reasons, it has been noted that a general trend in preparing textual data for analysis has been gradually to reduce or eliminate reliance on stop lists. Another approach to restricting the focus of textual data analysis from all words to only potentially informative words is to filter words by indices constructed from their relative frequency across as well as within documents. For such purposes, the "tf.idf" or "term frequency-inverse document" frequency measure (and its many variants) has become a staple of computational linguistics, usually serving as a weight to assign to words rather than as a simple inclusion or exclusion mechanism. The idea behind using tf.idf to weight textual data used in analysis is that the most informative words about a particular text are those that appear many times in that text but in relatively few others. At the extreme, weighting words by their tf.idf will exclude words that occur equally in all documents.

Other methods of preprocessing textual data aimed at generating units of analysis from sampling units include converting text to "*n*-grams," defined as sequences of *n* consecutive items, usually words in political science applications—thereby distinguishing "command economy" from "market economy." In an application of a political dictionary to text fragments where election manifestos defined the sampling units, for instance, one might define sampling units as 10-word sequences, irrespective of punctuation. Words may also be tagged by their part of speech, using language-specific algorithms developed in computational linguistics, with different parts of speech treated differently in whatever analysis follows.

Identifying units of analysis may also be done qualitatively, based on reading the texts and identifying politically relevant units of text. The best known example in political research is the "quasi sentence" that forms the unit of analysis for the long-running Comparative Manifesto Project. Quasi sentences are textual units that express a policy proposition and may be either a complete natural sentence or part of one. Because some authors may express two distinct policy statements within a single natural sentence, the use of "quasi" sentences permits a more valid and complete representation of the content of the textual data. The trade-off, however, is that the same human decision process that interprets the sentence structure to identify text units causes the procedure to be unreliable and often difficult or impossible to replicate. This trade-off between reliability—whether repetition of a procedure produces stable results—and validity—whether the measurement or analysis reflects the truth of what is being measured or represented by the textual data—is a recurrent theme in research involving textual data. This affects not just the identification and preparation of units for analysis but also the design of coding frames and measurement and scaling models for analyzing textual data.

Confronting these practical issues may seem like a potentially complex matter, but fortunately, for nearly every practical need, there exist several ready-made software solutions. Stemmers, stop word lists, word frequency generators, *n*-gram parsers, parts of speech identifiers, encoding converters, and a variety of utilities to translate from one format to another are readily, and often freely, available. The practical challenge of dealing with textual data then becomes more one of work flow management and research design than of the technical matter of solving specific practical needs.

Analyzing Textual Data

Once texts have been identified, selected, and prepared, they are ready for analysis. The analysis of textual data is not different from the analysis of any other data: no matter what its form, the analysis procedure is only of value if it can somehow produce propositions, measures, scales, or interpretations of lesser complexity than the textual data themselves. Analyzing textual data involves applying some explicit procedure to units of analysis from the texts, whether this procedure is purely qualitative, purely quantitative, or some hybrid in between. In practice, the variety of ways that text may be analyzed is vast. Some techniques are purely qualitative, involving interpretation of textual data by simply reading them and producing some summary account. Other qualitative methods involve reading a text and classifying the text according to a relatively simple coding frame, such as prowar, neutral, or antiwar. Yet other qualitative methods rely on human judgment to apply rules to units of textual data and possibly also in identifying those units. The Comparative Manifesto Project is an example of such a scheme, using human coders to identify and classify variable textual units using a 56-fold scheme of political policy categories: so is the 65-category coding scheme developed by Frank Baumgartner, Suzanna DeBoef, and Amber Boydstun in 2008 to capture arguments for and against the death penalty.

Many methods for analyzing textual data combine schemes involving human decisions and qualitative judgments with the use of computers to perform the mechanics of the processing and analysis. Tagging text with a predefined scheme or dictionary is an example of such a procedure. Several scaling algorithms, for example the "wordscores" method developed by Michael Laver, Kenneth Benoit, and John Garry, require human judgment to identify reference texts and to specify their relative locations on the scale of interest in order to estimate scaled positions for additional texts subsequently, using a purely automated procedure requiring no further human judgment.

The ultimate quantitative extreme in textual data analysis uses scaling procedures borrowed from item response theory methods developed originally in psychometrics. Both Jon Slapin and Sven-Oliver Proksch's Poisson scaling model and Burt Monroe and Ko Maeda's similar scaling method assume that word frequencies are generated by a probabilistic function driven by the author's position on some latent scale of interest and can be used to estimate those latent positions relative to the positions of other texts. Such methods may be applied to word frequency matrixes constructed from texts with no human decision making of any kind. The disadvantage is that while the scaled estimates resulting from the procedure represent relative differences between texts, they must be interpreted if a researcher is to understand what politically significant differences the scaled results represent. This interpretation is not always self-evident.

Recent textual data analysis methods used in political science have also focused on classification: determining which category a given text belongs to. Recent examples include methods to categorize the topics debated in the U.S. Congress as a means of measuring political agendas. Variants on classification include recently developed methods designed to estimate accurately the proportions of categories of opinions about the U.S. presidency from blog postings, even though the classifier on which it is based performs poorly for individual texts. New methodologies for drawing more information from political texts continue to be developed, using clustering methodologies, more advanced item response theory models, support vector machines, and semisupervised and unsupervised machine learning techniques.

Kenneth Benoit
Trinity College
Dublin, Ireland

See also Data, Archival; Discourse Analysis; Interviews, Expert; Party Manifesto

Further Readings

Baumgartner, F. R., DeBoef, S. L., & Boydstun, A. E. (2008). *The decline of the death penalty and the discovery of innocence.* Cambridge, UK: Cambridge University Press.

Klingemann, H.-D., Volkens, A., Bara, J., Budge, I., & McDonald, M. (2006). *Mapping policy preferences II: Estimates for parties, electors, and governments in Eastern Europe, European Union and OECD 1990–2003.* Oxford, UK: Oxford University Press.

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review, 97*(2), 311–331.

Leites, N., Bernaut, E., & Garthoff, R. L. (1951). Politburo images of Stalin. *World Politics, 3,* 317–339.

Monroe, B., & Maeda, K. (2004). *Talk's cheap: Text-based estimation of rhetorical ideal-points* (Working Paper). Lansing: Michigan State University.

Slapin, J. B., & Proksch, S.-O. (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science, 52*(3), 705–722.

Data Analysis, Exploratory

John W. Tukey, the definer of the phrase *exploratory data analysis* (EDA), made remarkable contributions to the physical and social sciences. In the matter of data analysis, his groundbreaking contributions included the fast Fourier transform algorithm and EDA. He reenergized descriptive statistics through EDA and changed the language and paradigm of statistics in doing so. Interestingly, it is hard, if not impossible, to find a precise definition of EDA in Tukey's writings. This is no great surprise, because he liked to work with vague concepts, things that could be made precise in several ways. It seems that he introduced EDA by describing its characteristics and creating novel tools. His descriptions include the following:

1. "Three of the main strategies of data analysis are: 1. graphical presentation. 2. provision of flexibility in viewpoint and in facilities, 3. intensive search for parsimony and simplicity." (Jones, 1986, Vol. IV, p. 558)
2. "In exploratory data analysis there can be no substitute for flexibility; for adapting what is calculated—and what we hope plotted—both to the needs of the situation and the clues that the data have already provided." (p. 736)
3. "I would like to convince you that the histogram is old-fashioned. . . ." (p. 741)

4. "Exploratory data analysis . . . does not need probability, significance or confidence." (p. 794)
5. "I hope that I have shown that exploratory data analysis is actively incisive rather than passively descriptive, with real emphasis on the discovery of the unexpected." (p. lxii)
6. "'Exploratory data analysis' is an attitude, a state of flexibility, a willingness to look for those things that we believe are not there, as well as those we believe to be there." (p. 806)
7. "Exploratory data analysis isolates patterns and features of the data and reveals these forcefully to the analyst." (Hoaglin, Mosteller, & Tukey, 1983, p. 1)
8. "If we need a short suggestion of what exploratory data analysis is, I would suggest that: 1. it is an attitude, AND 2. a flexibility, AND 3. some graph paper (or transparencies, or both)." (Jones, 1986, Vol. IV, p. 815)

This entry presents a selection of EDA techniques including tables, five-number summaries, stem-and-leaf displays, scatterplot matrices, box plots, residual plots, outliers, bag plots, smoothers, reexpressions, and median polishing. Graphics are a common theme. These are tools for looking in the data for structure, or for the lack of it.

Some of these tools of EDA will be illustrated here employing U.S. presidential elections data from 1952 through 2008. Specifically, Table 1 displays the percentage of the vote that the Democrats received in the states of California, Oregon, and Washington in those years. The percentages for the Republican and third-party candidates are not a present concern. In EDA, one seeks displays and quantities that provide insights, understanding, and surprises.

Table

A table is the simplest EDA object. It simply arranges the data in a convenient form. Table 1 is a two-way table.

Five-Number Summary

Given a batch of numbers, the five-number summary consists of the largest, smallest, median, and upper and lower quartiles. These numbers are useful for auditing a data set and for getting a feel for the data. More complex EDA tools may be based on them. For the California data, the five-number summary in percents is shown in Figure 1.

These data are centered at 47.6% and have a spread measured by the interquartile range of 8.75%. Tukey actually employed related quantities in a hope to avoid confusion.

Stem-and-Leaf Display

The numbers of Table 1 provide all the information, yet condensations can prove better. Figure 2

Table 1 Percentages of the Votes Cast for the Democratic Candidate in the Presidential Years 1952–2008

	Year														
State	*1952*	*1956*	*1960*	*1964*	*1968*	*1972*	*1976*	*1980*	*1984*	*1988*	*1992*	*1996*	*2000*	*2004*	*2008*
California	42.7	44.3	49.6	59.1	44.7	41.5	47.6	35.9	41.3	47.6	46.0	51.1	53.4	54.3	61.0
Oregon	38.9	44.8	44.7	63.7	43.8	42.3	47.6	38.7	43.7	51.3	42.5	47.2	47.0	51.3	56.7
Washington	44.7	45.4	45.4	62.0	47.2	38.6	46.1	37.3	42.8	50.1	45.1	49.8	50.2	52.8	57.7

Source: Statistical Abstracts of the U.S. Census Bureau.

Minimum	*Lower Quartile*	*Median*	*Upper Quartile*	*Maximum*
35.9	43.50	47.6	52.25	61.0

Figure 1 A Five-Number Summary for the California Democrat Percentages

Note: The minimum, 35.9%, occurred in 1980 and the maximum, 61%, in 2008.

3 I 6	3 I 6
4 I 12345688	4 I 1234
5 I 01349	4 I 5688
6 I 1	5 I 0134
	5 I 9
	6 I 1

Figure 2 Stem-and-Leaf Displays, With Scales of 1 and 2, for the California Democratic Data

provides a stem-and-leaf display for the data of the table. There are stems and leafs. The stem is a line with a value. See the numbers to the left of the " I ". The leaves are numbers on a stem, the right-hand parts of the values displayed.

Using this exhibit, one can read, off various quartiles, the five-number summary approximately; see indications of skewness; and infer multiple modes.

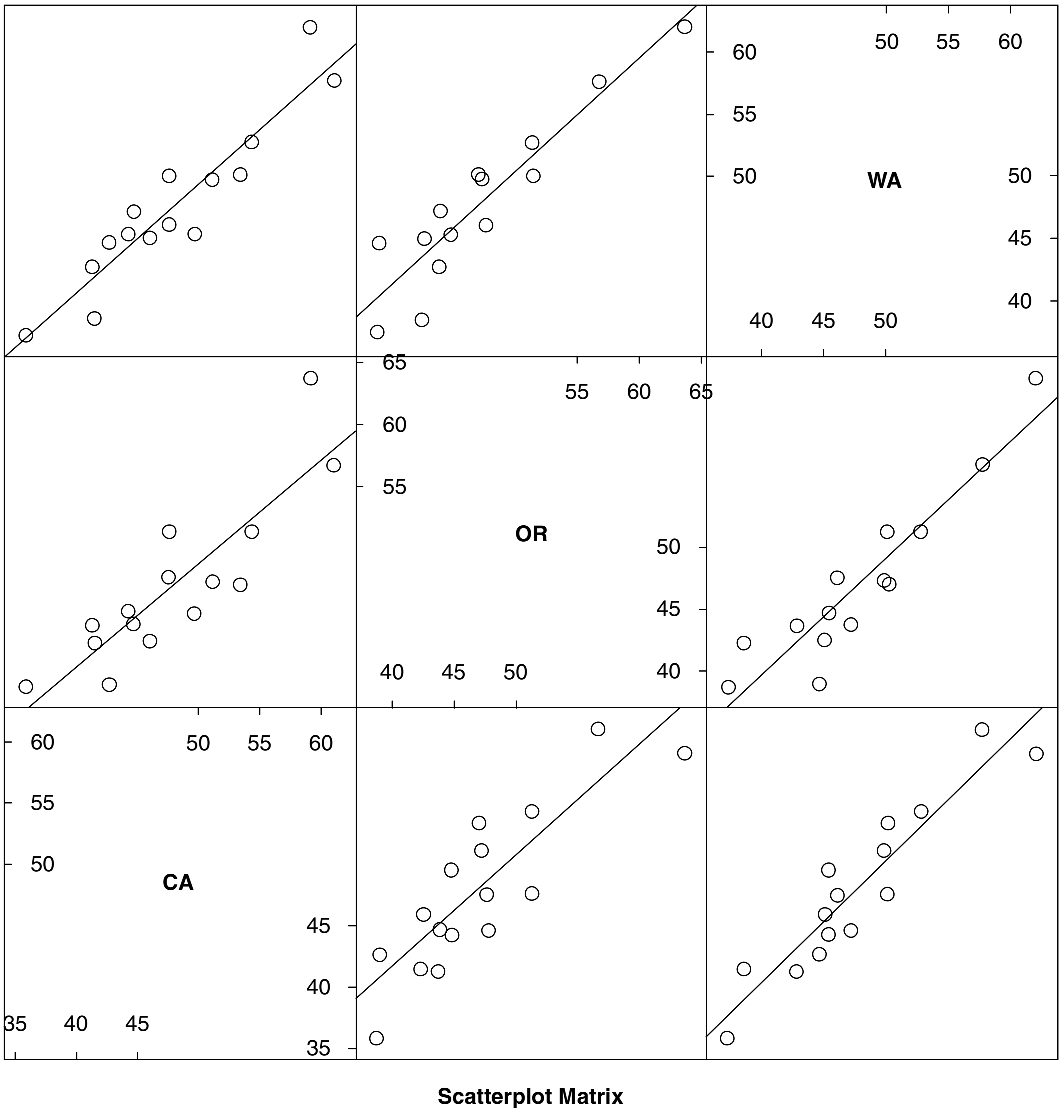


Figure 3 Scatterplots of Percentages for the States Versus Percentages for the States in Pairs

Notes: A least squares line has been added as a reference. CA = California; OR = Oregon; WA = Washington.

Scatterplot Matrix

Figure 3 displays individual scatterplots for the state pairs (CA, OR), (CA, WA), and (OR, WA). A least squares line has been added in each display to provide a reference. One sees the x and y values staying together. An advantage of the figure over three individual scatterplots is that one sees the plots simultaneously.

Outliers

An outlier is an observation strikingly far from some central value. It is an unusual value relative to the bulk of the data. Commonly computed quantities such as averages and least squares lines can be drastically affected by such values. Methods to detect outliers and to moderate their effects are needed. So far, the tools discussed in this entry have not found any clear outliers.

Box Plots A box plot consists of a rectangle with top and bottom sides at the levels of the quartiles, a horizontal line added at the level of the median, and whiskers, of length 1.5 times the interquartile range, added at the top and bottom. It is based on numerical values. Points outside these limits are plotted and are possible outliers. Figure 4 presents three box plots. When more than one box plot are present in a figure, they are referred to as *parallel box plots*.

Figure 4 presents a parallel box plot display for the presidential data. The California values tend to be higher than those of Oregon and Washington. Those show a single outlier each and a skewing toward higher values. Both the outliers are for the 1964 election.

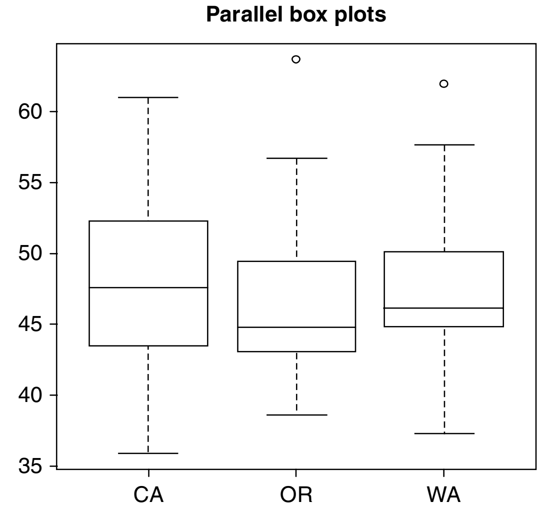


Figure 4 Parallel Box Plots for the Percentages, One for Each State

Note: CA = California; OR = Oregon; WA = Washington.

Residual Plots

A residual plot is another tool for detecting outliers and noticing unusual patterns. Suppose there is a fit to the data, say, a least squares line. The residuals are then the differences between the data and their corresponding fitted values.

Consider the percentages in the table depending on the year of the election—that is, consider the data as a time series (Figure 5).

The time series of these three states track each other very well, and there is a suggestion of an outlier in each plot.

Figure 6 shows the residuals for the three states.

Each display in Figure 6 shows an outlier near the top. They all correspond to year 1964. This was the first year after John Kennedy was assassinated, and Lyndon Johnson received a substantial sympathy vote. There also is a suggestion of temporal dependence.

With today's large data sets, one wishes for automatic ways to identify and handle outliers and other unusual values. One speaks of resistant/robust methods, resistant methods being those not overly sensitive to the presence of outliers and robust ones being those not affected strongly by long tails in the distribution. In the case of bivariate data, one can consider the bag plot.

Bag Plots

The bag plot is a generalization of the box plots of Figure 4. It is often a convenient way to study the scatter of bivariate data. In the construction of a bag plot, one needs a bivariate median, analogs of the quartiles, and whiskers. Tukey and his collaborators developed these. The center of the bag plot is the Tukey median. The "bag" surrounds the center and contains the 50% of the observations with the greatest depth. The "fence" separates inliers from

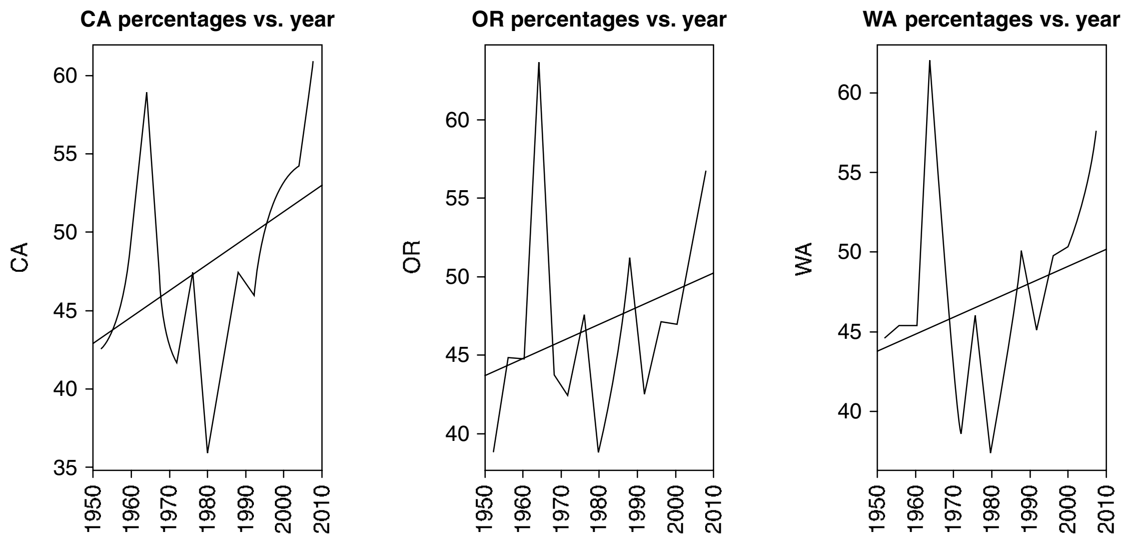


Figure 5 Graphs of the Individual State Democrat Percentages Versus the Election Years

Notes: A least squares line has been added as a reference. CA = California; OR = Oregon; WA = Washington.

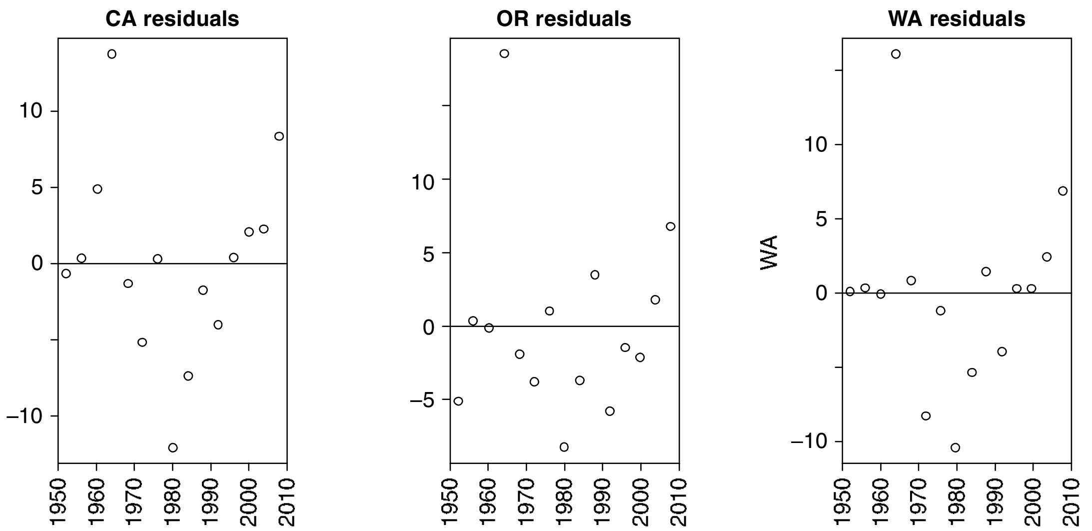


Figure 6 Residuals From Least Squares Line Versus Year With 0-Line Added

Note: CA = California; OR = Oregon; WA = Washington.

outliers. Lines called whiskers mark observations between the bag and the fence. The fence is obtained by inflating the bag, from the center, by a factor of 3.

Figure 7 provides bag plots for each of the pairs (CA, OR), (CA, WA), and (OR, WA).

One sees an apparent outlier in both the California versus Oregon and the California versus Washington cases. Interestingly there is not one for the Oregon versus Washington case. On inspection, it is seen that the outliers correspond to the 1964 election. One also sees that the points

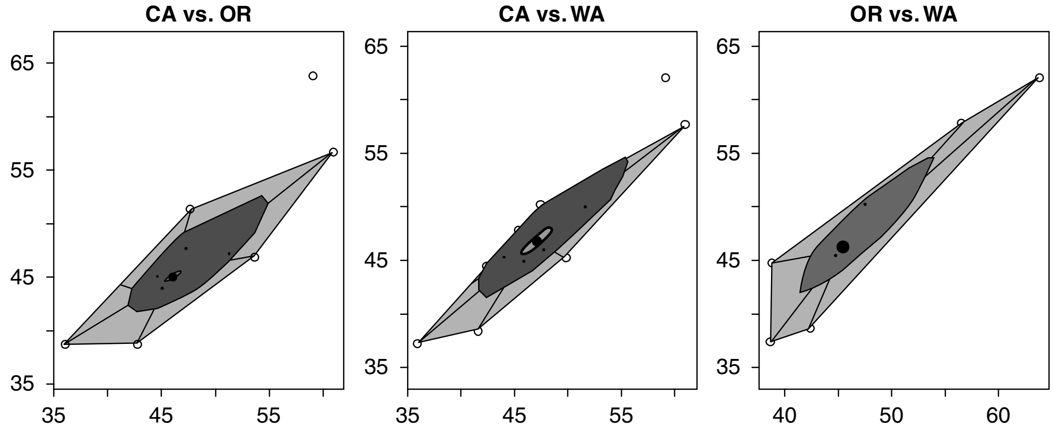


Figure 7 Bag Plots for the State Pairs, Percentages Versus Percentages

Note: The bivariate median is the black spot within the bag (darker shading), and the fence is the outer boundary.

vaguely surround a line. Because of the bag plot's resistance to outliers, the unusual point does not affect its location and shape.

Smoother

Smoothers have as a goal the replacement of a scatter of points by a smooth curve. Sometimes the effect of smoothing is dramatic and a signal appears. The curve resulting from smoothing might be a straight line. More usefully, a local least squares fit might be employed with the local curves, $y = f(x)$, a quadratics. The local character is often introduced by employing a kernel. A second kernel might be introduced to make the operation robust/resistant. It will have the effect of reducing the impact of points with large residuals.

Figure 8 shows the result of local smoothing of the Democrat percentages as a function of election year. The loess procedure was employed.

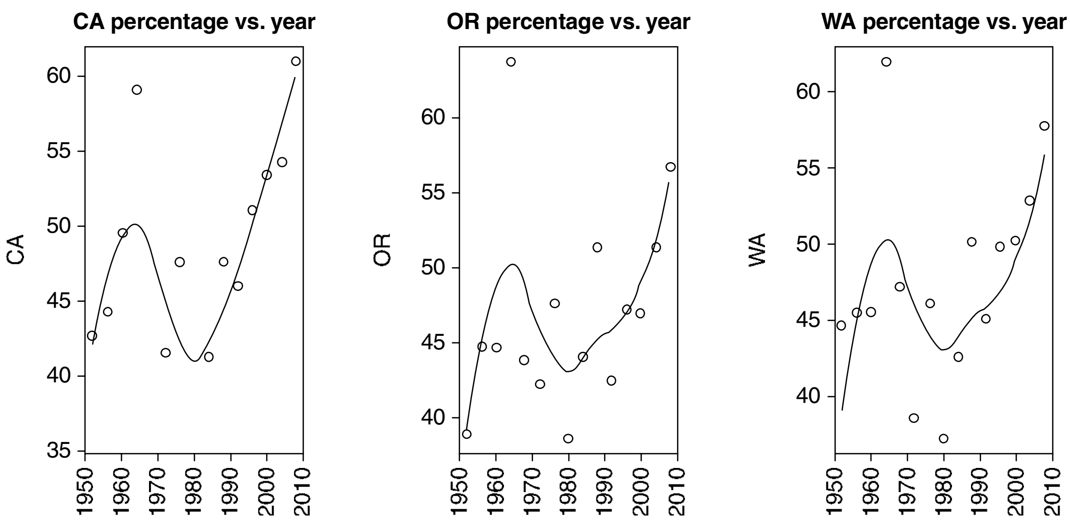


Figure 8 Percents Versus Year With a Loess Curve Superposed

Note: CA = California; OR = Oregon; WA = Washington.

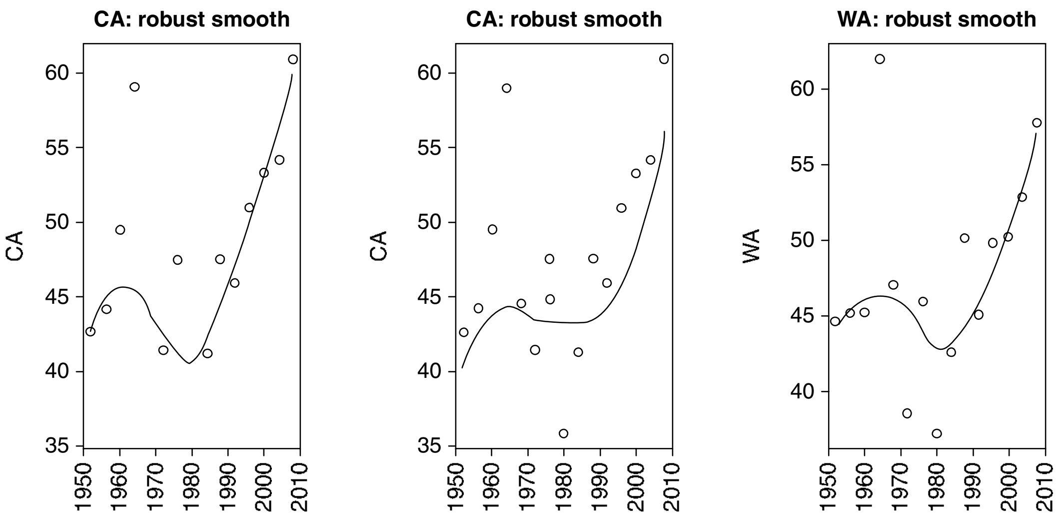


Figure 9 The Curves Plotted Are Now Resistant to Outliers

Note: CA = California; OR = Oregon; WA = Washington.

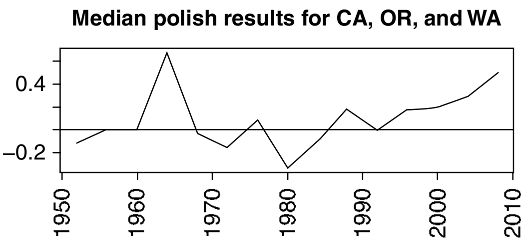


Figure 10 The Year Effect Obtained for the Election Data via Median Polishing

Note: CA = California; OR = Oregon; WA = Washington.

The curves have a similar general appearance. They are pulled up by the outlier at the 1964 point.

Robust Variant

The behavior in 1964 being understood to a degree, one would like an automatic way to obtain a curve not so strongly affected by this outlier. The loess procedure has a robust/resistant variant. The results follow in Figure 9.

Having understood that 1964 was an unusual year, one can use a robust curve to understand the other values better. The plots have similar shapes. One sees a general growth in the Democrat percentages starting around 1980. In this two-step procedure, it is important to study both the outlier and the robust/resistant curve.

Reexpression

This term refers to expressing the same information by different numbers, for example, using logit = $\log(p/(1 - p))$ instead of the proportion p. The purpose may be additivity, obtaining straightness or symmetry, or making variability more nearly uniform.

The final method is a tool for working with two-way tables.

Median Polish

This is a process of alternately finding and subtracting medians from rows and then columns and perhaps continuing to do this until the results do not change much. One purpose is to seek an additive model for a two-way table, in the presence of outliers in the data.

The state percentages in Table 1 form a 3×15 table and a candidate for median polish. The resulting row (year) effects are shown in Figure 10.

These effects are not meant to be strongly affected by outliers. Figure 10 shows the same general curve as in Figure 5.

This entry ends with Tukey's 1973 rejoinder: "Undoubtedly, the swing to exploratory data analysis will go somewhat too far" (cited in Jones, Vol. III, p. lxii).

David R. Brillinger
University of California, Berkeley
Berkeley, California, United States

See also Cross-Tabular Analysis; Data Visualization; Graphics, Statistical; Statistics: Overview

Further Readings

Bashford, K. E., & Tukey, J. W. (1999). *Graphical analysis of multivariate data.* London: Chapman & Hall.

Brillinger, D. R. (2002). John W. Tukey: The life and professional contributions. *Annals of Statistics, 30,* 1535–1575.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis.* New York: Wiley.

Jones, L. V. (Ed.). (1986). *The collected works of John W. Tukey: Philosophy and principles of data analysis 1949–1964* (Vols. III & IV). London: Chapman & Hall.

McNeil, D. R. (1977). *Interactive data analysis.* New York: Wiley.

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics, 33,* 1–67.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, PA: Addison-Wesley.

Velleman, P. F., & Hoaglin, D. C. (1981). *Applications, basics, and computing of exploratory data analysis.* Boston: Duxbury.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer.

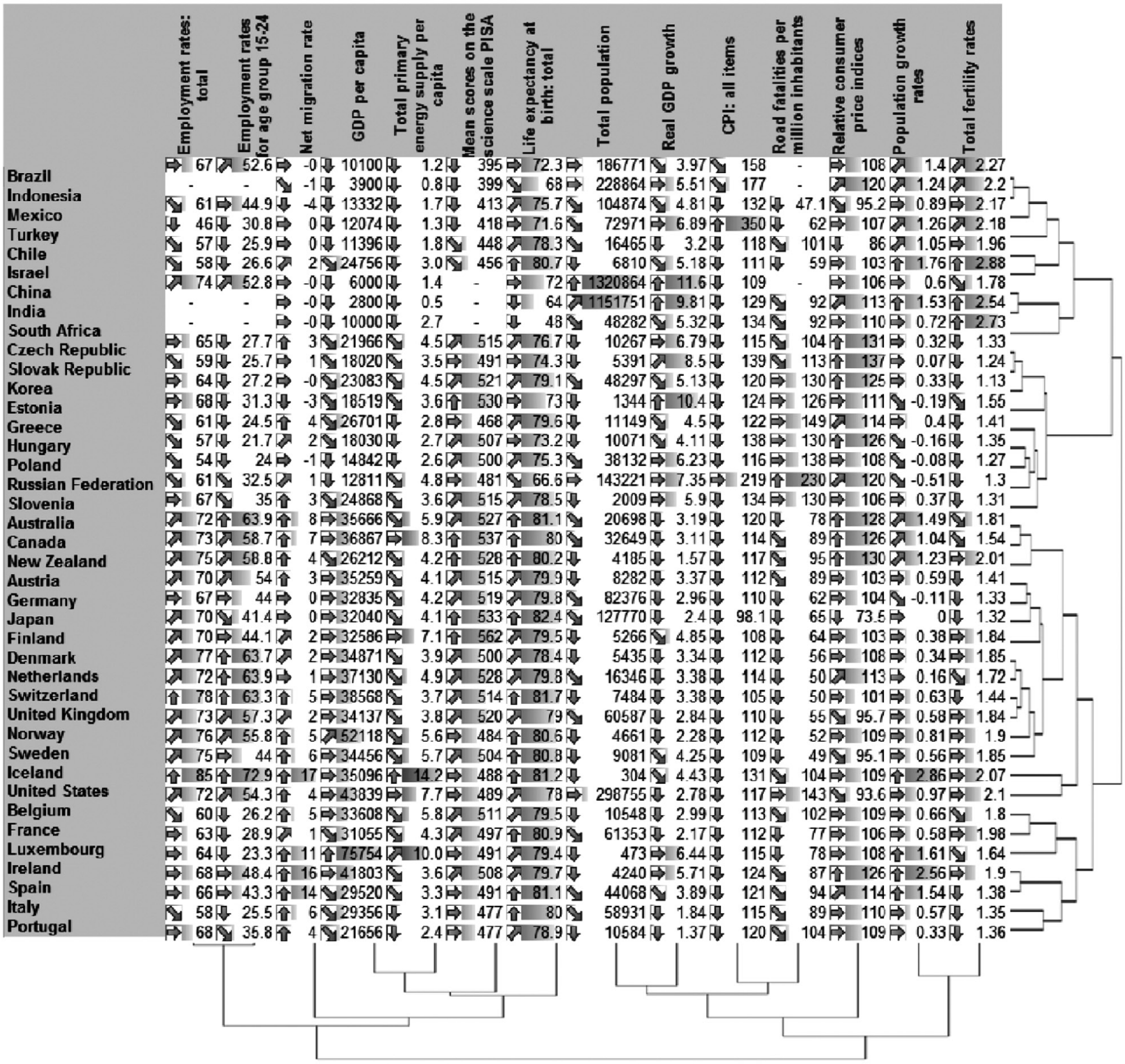
Data Visualization

The basic objective of data visualization is to provide an efficient graphical display for summarizing and reasoning about quantitative information. Data visualization should be distinguished from other types of visualization used in political science (more general information and knowledge visualization, concept visualization, strategy and work flow visualization, metaphor visualization, etc.) as it is more specific to the representation of quantitative data existing in the form of numerical tables. In the following sections, the different types and methods of data visualization and their application in political science are presented.

Chart Types and Methods

During the past decades, political science has accumulated a large corpus of various kinds of data such as comprehensive fact books and atlases, characterizing all or most of existing states by multiple and objectively assessed numerical indicators within certain time periods (e.g., *OECD Factbook* and *Political Atlas of the Modern World*). As a consequence, there exists a tendency for political science to gradually become a more quantitative scientific field and to use quantitative information in analysis and reasoning. Any analysis in political science must be multidimensional and combine various sources of information; however, human capabilities for perception of large amounts of numerical information are limited. Hence, methods and approaches for the visualization of quantitative and qualitative data (especially multivariate data) are an extremely important topic in political science. Data visualization approaches can be classified into several groups, starting from creating informative charts and diagrams (statistical graphics and infographics) and ending with advanced statistical methods for visualizing multidimensional tables containing both quantitative and qualitative information. Data visualization in political science takes advantage of recent developments in computer science and computer graphics, statistical methods, methods of information visualization, visual design, and psychology. Data visualization in political science has certain special features such as the frequent use of geographical maps, which creates a link with the well-developed field of geographic information systems (GIS). Furthermore, numerical tables in political science are often incomplete, which makes important the use of methods dealing with missing or uncertainly measured data entries.

There are two main types of numerical tables that can be the subject of data visualization. The first one is called an *object–feature table*, where each row represents an observation or an object and each column corresponds to a numerical feature or indicator commonly measured for the whole set of objects. An example of such an object–feature table



	Employment rates: total	Employment rates for age group 15-24	Net migration rate	GDP per capita	Total primary energy supply per capita	Mean scores on the science scale PISA	Life expectancy at birth: total	Total population	Real GDP growth	CPI: all items	Road fatalities per million inhabitants	Relative consumer price indices	Population growth rates	Total fertility rates
Brazil	67	52.6	-0	10100	1.2	395	72.3	186771	3.97	158	-	108	1.4	2.27
Indonesia	-	-	-1	3900	0.8	399	68	228864	5.51	177	-	120	1.24	2.2
Mexico	61	44.9	-4	13332	1.7	413	75.7	104874	4.81	132	47.1	95.2	0.89	2.17
Turkey	46	30.8	0	12074	1.3	418	71.6	72971	6.89	350	62	107	1.26	2.18
Chile	57	25.9	0	11396	1.8	448	78.3	16465	3.2	118	101	86	1.05	1.96
Israel	58	26.6	2	24756	3.0	456	80.7	6810	5.18	111	59	103	1.76	2.88
China	74	52.8	-0	6000	1.4	-	72	1320864	11.6	109	-	106	0.6	1.78
India	-	-	-0	2800	0.5	-	64	1151751	9.81	129	92	113	1.53	2.54
South Africa	-	-	-0	10000	2.7	-	48	48282	5.32	134	92	110	0.72	2.73
Czech Republic	65	27.7	3	21966	4.5	515	76.7	10267	6.79	115	104	131	0.32	1.33
Slovak Republic	59	25.7	1	18020	3.5	491	74.3	5391	8.5	139	113	137	0.07	1.24
Korea	64	27.2	-0	23083	4.5	521	79.1	48297	5.13	120	130	125	0.33	1.13
Estonia	68	31.3	-3	18519	3.6	530	73	1344	10.4	124	126	111	-0.19	1.55
Greece	61	24.5	4	26701	2.8	468	79.6	11149	4.5	122	149	114	0.4	1.41
Hungary	57	21.7	2	18030	2.7	507	73.2	10071	4.11	138	130	126	-0.16	1.35
Poland	54	24	-1	14842	2.6	500	75.3	38132	6.23	116	138	108	-0.08	1.27
Russian Federation	61	32.5	1	12811	4.8	481	66.6	143221	7.35	219	230	120	-0.51	1.3
Slovenia	67	35	3	24868	3.6	515	78.5	2009	5.9	134	130	106	0.37	1.31
Australia	72	63.9	8	35666	5.9	527	81.1	20698	3.19	120	78	128	1.49	1.81
Canada	73	58.7	7	36867	8.3	537	80	32649	3.11	114	89	126	1.04	1.54
New Zealand	75	58.8	4	26212	4.2	528	80.2	4185	1.57	117	95	130	1.23	2.01
Austria	70	54	3	35259	4.1	515	79.9	8282	3.37	112	89	103	0.59	1.41
Germany	67	44	0	32835	4.2	519	79.8	82376	2.96	110	62	104	-0.11	1.33
Japan	70	41.4	0	32040	4.1	533	82.4	127770	2.4	98.1	65	73.5	0	1.32
Finland	70	44.1	2	32586	7.1	562	79.5	5266	4.85	108	64	103	0.38	1.84
Denmark	77	63.7	2	34871	3.9	500	78.4	5435	3.34	112	56	108	0.34	1.85
Netherlands	72	63.9	1	37130	4.9	528	79.8	16346	3.38	114	50	113	0.16	1.72
Switzerland	78	63.3	5	38568	3.7	514	81.7	7484	3.38	105	50	101	0.63	1.44
United Kingdom	73	57.3	2	34137	3.8	520	79	60587	2.84	110	55	95.7	0.58	1.84
Norway	76	55.8	5	52118	5.6	484	80.6	4661	2.28	112	52	109	0.81	1.9
Sweden	75	44	6	34456	5.7	504	80.8	9081	4.25	109	49	95.1	0.56	1.85
Iceland	85	72.9	17	35096	14.2	488	81.2	304	4.43	131	104	109	2.86	2.07
United States	72	54.3	4	43839	7.7	489	78	298755	2.78	117	143	93.6	0.97	2.1
Belgium	60	26.2	5	33608	5.8	511	79.5	10548	2.99	113	102	109	0.66	1.8
France	63	28.9	1	31055	4.3	497	80.9	61353	2.17	112	77	106	0.58	1.98
Luxembourg	64	23.3	11	75754	10.0	491	79.4	473	6.44	115	78	108	1.61	1.64
Ireland	68	48.4	16	41803	3.6	508	79.7	4240	5.71	124	87	126	2.56	1.9
Spain	66	43.3	14	29520	3.3	491	81.1	44068	3.89	121	94	114	1.54	1.38
Italy	58	25.5	6	29356	3.1	477	80	58931	1.84	115	89	110	0.57	1.35
Portugal	68	35.8	4	21656	2.4	477	78.9	10584	1.37	120	104	109	0.33	1.36

Figure 1 Examples of Data Visualization for a Small Subset of Data (14 Socioeconomic Numerical Indicators for 40 Countries)

Source: Organisation for Economic Co-operation and Development. (2006). *OECD factbook 2006: Economic, environmental and social statistics*. Paris: Author.

Notes: The indicators describe education, energy consumption, employment situation, macroeconomic trends, population and migration, prices, and quality-of-life aspects of each country. The table itself is a graphical information display. The bars and arrows in the table cells show the indicator values (arrow up—above the median; arrow down—below the median; arrow right—near the median). The rows and columns are reordered so that the correlated indicators and statistical country profiles are ranked closely. The dendrogram on the right (or at the bottom) shows the separation of countries into groups (or grouping correlated indicators) as suggested by hierarchical clustering. CPI = Corruption Perception Index; GDP = gross domestic product.

is a fact book for a set of countries (see Figure 1), where the objects are countries and the features are numerical indicators such as gross domestic product (GDP) per capita, employment rate, life expectancy, and so on. The second type of numerical tables is called *connection or distance tables*, where both rows and columns correspond to objects and a numerical value characterizing a link between two objects is found at the intersection of a row and a column. Alternatively, such tables can be represented in the form of a list of links. A typical example of a connection table is the table representing migration rates or the mutual volumes of exports and imports for a set of countries (at the intersection of Row A and Column B, the volume of export from Country A to County B is found).

Data visualization plays several important roles in science: It (a) helps create informative illustrations of the data, recapitulating large amounts of quantitative information on a diagram; (b) helps formulate new hypotheses or to confirm existing hypotheses for quantitative data; and (c) guides a statistical analysis of data and checks its validity. One can also mention the role of infographics in creating images with a clear and visual message, based on numbers; thus, data visualization can serve as a powerful public relations or educational tool. Graphical display allows not only visualizing and analyzing the message contained in data but also remembering it, since for most people visual memory is more persistent than verbal or auditory memory (the phenomenon of pictorial superiority).

Several groups of methods for data visualization can be distinguished in political science:

- *statistical graphics and infographics*, with extensive use of color, form, size, shape, and style to superimpose many quantitative variables in the same chart or diagram;
- *geographic information systems* to visualize geographically linked data;
- *graph visualization* or *network maps* for representing the relations between objects; and
- *data cartography*—that is, the projection of multidimensional data on low-dimensional screens with further visualization.

As follows, a description of these principal groups of data visualization methods, along with references to examples of their application in political science, is provided.

Statistical Graphics and Infographics

Statistical graphics started to be used in science in the 18th century but began to be widely exploited only from the end of the 19th century. Many of the currently used types of charts were introduced by William Playfair, a Scottish political economist, and Johann Heinrich Lambert, a Swiss German mathematician. With the appearance of computer-based technologies toward the end of the 20th century, the use of statistical graphics exploded and included exploitation of three-dimensional, dynamic, and interactive data representations.

Ralph Lengler and Martin Eppler suggested a "periodic table" of visualization methods with many examples of elements of statistical graphics. The most commonly used quantitative data displays can be classified as follows:

- *Univariate plots* are designed to visualize value distribution of a single variable. The basic ones are pie charts, bar charts, histograms, box plots (see Figure 2).
- *Bivariate plots*, are designed to visualize relations between two variables. These include various types of scatterplots and line plots (see Figure 3). A particularly important case is the line plot–visualizing time series in which one variable is time.
- *Multivariate plots*, designed to visualize the values of several variables at the same time and comparing them. Examples of these are area charts, radar charts, mosaic plots, parallel coordinates plots, and the more sophisticated Chernoff faces.

This classification is elementary; however, modern statistical graphics usually superimpose several types of plots and use color, shape, size, and style of their elements to increase the dimensionality of the plot. New types of data displays are invented continuously and proposed as parts of commonly used software for data analysis. As an example, the data table itself can serve as a data display (see Figure 1). Moreover, the table can be converted into the more elaborate information lens display. Time-series representation can be significantly enhanced by colorful horizon graphs that allow visualizing hundreds of time series simultaneously and identifying patterns in their behavior visually. Treemaps are a space-filling approach to show hierarchies in which the rectangular screen space is divided into regions, and then each region is divided again for each level in the hierarchy (e.g., treemaps were used for visualization of the U.S. budget). Intersections

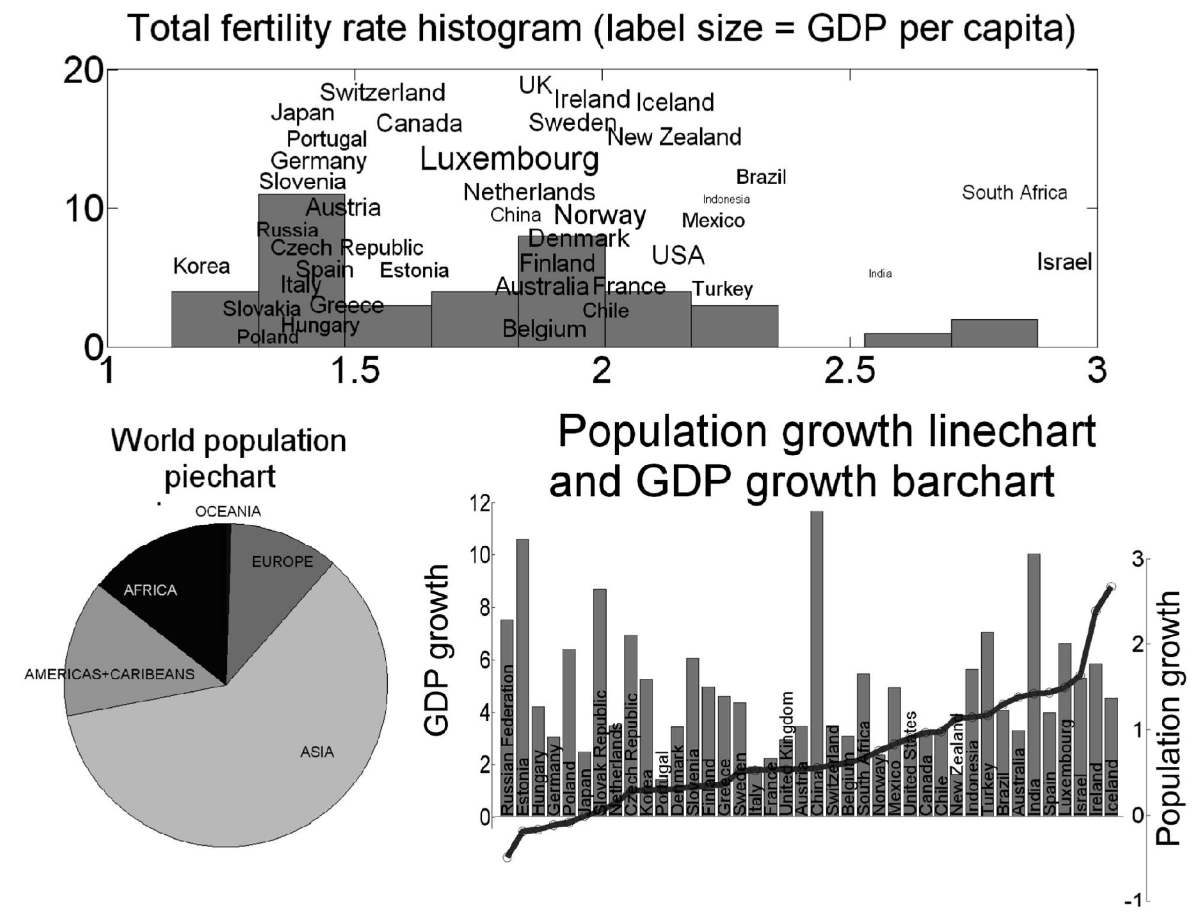


Figure 2 Examples of Standard Statistical Graphs

Source: Organisation for Economic Co-operation and Development. (2006). *OECD factbook 2006: Economic, environmental and social statistics*. Paris: Author.

between sets of objects can be represented by Venn diagrams. Many useful graphical displays result from application of multidimensional statistical analysis techniques such as factor analysis (including principal components analysis) and clustering (see Figures 1 and 4 and the section on multidimensional data cartography).

In the physical, biological, and social sciences, the scatterplot is the predominant type of data display (according to Edward R. Tufte, approximately 75% of graphs used in science are scatterplots). Scatterplots serve two basic purposes: (1) to visually detect linear and nonlinear relations between two variables; the human eye can do this efficiently and is extremely robust to the effects of anomalous observations (outliers) and other aberrations in the data, and (2) to create a map of data allowing one to visualize other information on top of it. In the case of a large number of points, the scatterplot can be improved by visualizing the point density, using gradient shading (Figure 4) or isodensity levels (i.e., contours connecting points with the same density).

In contrast, it was shown that the use of statistical graphics in news media is mainly limited to univariate plots (if time-series plots are not considered), mainly bar charts and pie charts. Curiously, several Japanese news journals and the British *Economist* create notable exceptions and provide up to 9% of data visualizations in the form of bivariate plots. One of the first books on the history of data visualization, *Graphic Presentation,* was written by Willard Brinton in 1939. It contains

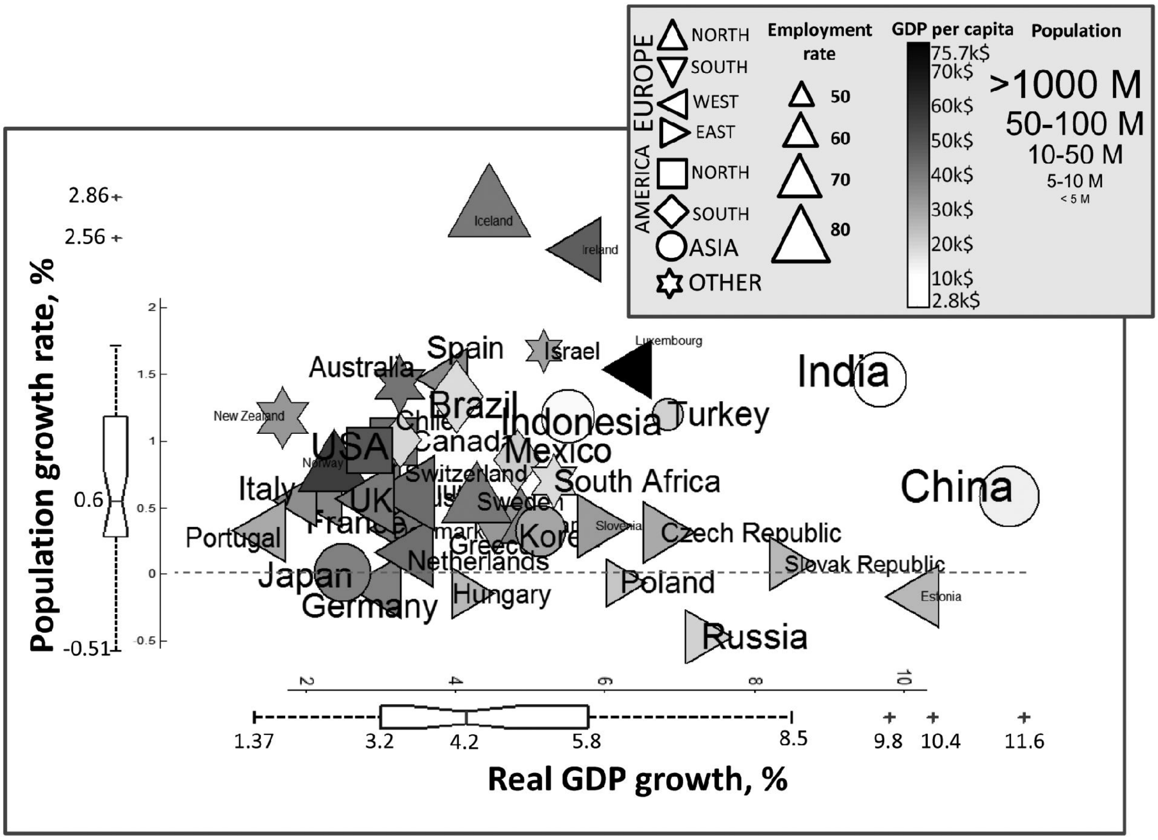


Figure 3 Scatterplot as a Means to Visualize the Values of Five Indicators Simultaneously

Source: Organisation for Economic Co-operation and Development. (2006). *OECD factbook 2006: Economic, environmental and social statistics*. Paris: Author.

Notes: The five indicators are GDP and population growth rates, using abscissa and ordinate point position; employment rate, using marker size; GDP per capita, using marker color; and population, using the text size of the data labels. In addition, the geographical localization is visualized in the form of the marker. The box plots on the axes depict the median value, lower and upper quartiles, extent of the data, and outlier values (by crosses). GDP = gross domestic product.

many examples of thoughtful data visualization in the precomputer era.

Geographic Information Systems

In 1854, John Snow depicted a cholera outbreak in London and detected the source of the epidemic using points to localize some individual cases. This was one of the earliest applications of the geographic method—that is, combining the use of a geographical map and statistical data. Nowadays, the use of geographical maps with several layers of information (also called thematic maps or semantic layers) is facilitated by GIS. GIS is the merging of cartography, statistical analysis, and database technology. Geovisualization tools have been used extensively in electoral studies, urbanization studies, and various geopolitical studies of empires, wars, boundaries, and trade routes.

In the mid-2000s, several publicly available GIS systems called virtual globes appeared, with Google Earth being the most popular one. These systems allow the visualization of geographical maps and various types of global and local semantic layers mapped on them in a highly dynamic, interactive fashion. Importantly, these systems themselves can collect information from users (Public Participation GIS). Because of their wide use by the public for

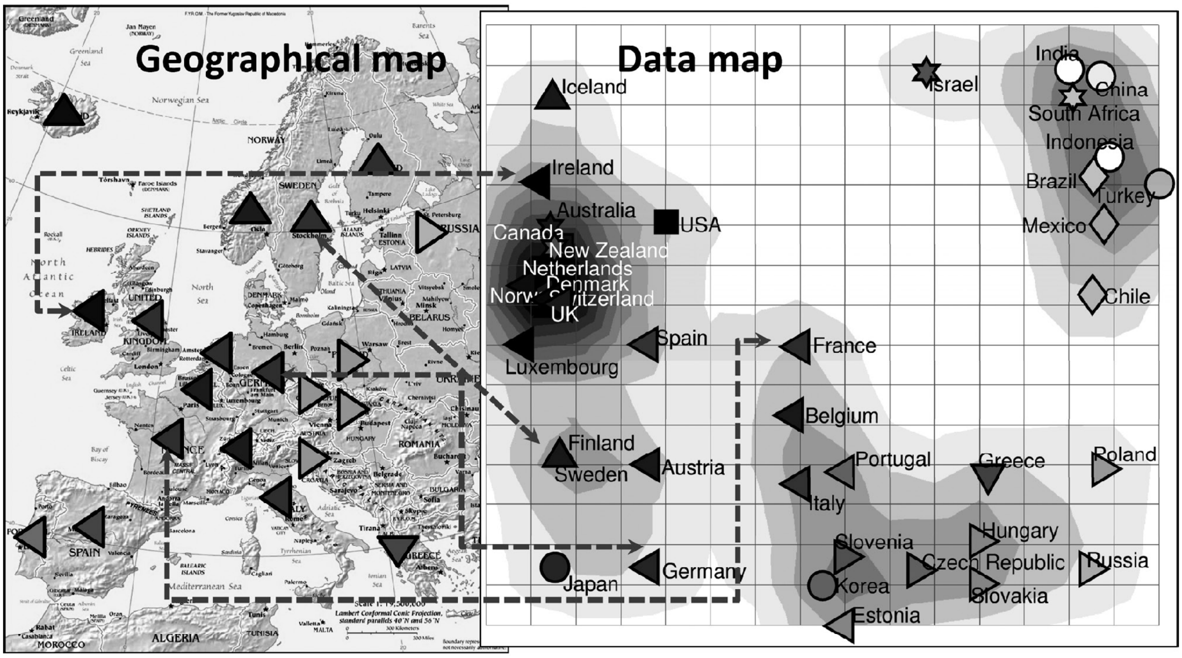


Figure 4 Data Cartography for Multivariate Data Visualization

Source: Organisation for Economic Co-operation and Development. (2006). *OECD factbook 2006: Economic, environmental and social statistics*. Paris: Author.

Notes: The nonlinear data map is shown on the right and created with the use of the elastic map algorithm implemented in ViDaExpert software (http://bioinfo.curie.fr/projects/vidaexpert). The gray shading represents the density of points (from this semantic map, four groups of countries can be distinguished). The geographical map on the left shows geographical localization of some of the points corresponding to European countries. The marker gray scale here represents GDP per capita (from white to black). The countries close on the geographical map are colocalized physically, while the countries close on the data map are neighbors in the multidimensional space of indicators and have similar statistical profiles. The axes of the data map themselves do not have an explicit interpretation; what matters is the distance between points, showing similar statistical profiles for points with a short distance and different ones for those with a larger distance. GDP = gross domestic product.

various purposes, these systems have a high potential to be used in political and social studies and are the subject of several educational courses in social science.

Graph (Network) Visualization

Graph visualization (where a graph is understood as a set of vertices connected by arrows, directed or undirected), also known as network maps, is a graphical data display that allows one to visualize a specific type of connection tables. On such a display, a vertex represents an object or an observation, and an arrow between two vertices represents a connection between two objects. Usually, only arrows representing nonempty connections (or having a strength exceeding a certain threshold) are introduced into the graph. On the network maps, one can use the size and the color of vertices to visualize some features of the objects and the thickness and the color of arrows to represent some features of connections between objects (typically, the strength of the connection). If the connections are asymmetrical (e.g., export and import flows, migration rates between countries), then two directed arrows connecting the same objects can be used to represent the asymmetric flow in both directions. Network maps can also represent distances between objects in the multidimensional space of multivariate observations. Network maps are widely used for visualizing social network structures, trading relations between countries, migration rates, and so on. A good example of such visualization is the Mapping Globalization Project of Princeton University.

The principal difficulty in using network map displays is that they easily become too entangled with a growing number of arrows. This poses technical challenges for (a) finding a good layout for placing vertices on a two-dimensional plane to minimize the number of arrow intersections or visualize the internal structure of the graph better and (b) creating handy browsers for huge network presentations. For example, in the Walrus network visualization tool, three-dimensional interactive network map representations are combined with a hyperbolic viewer that creates a fish eye–like distortion to magnify (to zoom) a particular part of the graph while the rest of it remains less detailed and serves rather for orientation and navigation.

Multidimensional Data Cartography

Inventors of statistical graphics thought of data displays as analogues of geographical maps. For bivariate data visualization plots, it was a crucial invention to replace geographical latitude and longitude with arbitrary measurement axes, which required the notion of a coordinate system formalized in René Descartes' *La Géometrie* written in 1637. Thus, on a scatterplot, instead of putting geographically colocated objects together, one puts together the points corresponding to similar combinations of measured x and y values.

For multivariate observations, one can formulate a similar problem: How to map a set of objects (vectors) from a multidimensional space onto a two-dimensional plane (or into a three-dimensional space) such that the objects with similar numerical feature profiles would be located close and dissimilar objects would be located at a larger distance after projection. There are a number of statistical methods aiming at producing such data maps. That is a part of exploratory data analysis.

The most fundamental method for mapping multidimensional data into low-dimensional space is the principal component analysis (PCA), invented by Karl Pearson. The method can be applied for the numerical tables of the object–feature type. The set of table rows is represented as a cloud of data points (vectors) in the multidimensional space of features. Then the PCA constructs an optimal linear two-dimensional screen (a plane) embedded into the multidimensional space of data such that the sum of squared distances between the data points and the screen is minimal. After that, the data points can be projected onto the screen and the distribution of the projections is represented using standard statistical graphics techniques. The great advantage of such visualization is that it takes into account all numerical dimensions at the same time and not only two dimensions as on the standard bivariate scatterplots. The disadvantage of all multidimensional data mapping methods is that the new axes of the resulting scatterplot do not have explicit meaning (in general, they are complex functions of the initial numerical features). What matters on such data visualization displays are the relative distances between projections that represent the distances between objects in the initial multidimensional space. The method of principal components as a multivariate data visualization tool was used in several studies, including *Political Atlas of the Modern World*. The classical PCA method can have nonlinear generalizations (principal manifolds approach), and in this case, it becomes more precise and informative in visualizing the data structure. In this case, a nonlinear (curved) two-dimensional screen is constructed in the multidimensional space and used in the same way for projecting the data as in classical PCA. Other multivariate data-mapping techniques such as correspondence analysis, self-organizing maps, metro maps visualization of principal trees, multidimensional scaling, locally linear embedding, and ISOMAP have been recently developed and applied in political and social science.

Visualizing multivariate data by projection onto a low-dimensional screen is a subject of multidimensional data cartography: The projection creates a data map that is an alternative to the geographical map (see Figure 4) and on which the objects (not necessarily linked to geography) with similar feature profiles are colocalized. For these data maps, the methods of data visualization developed for GIS can be reused. For example, an atlas of semantic layers can be created, each layer corresponding to the values of a particular numerical attribute.

Data Visualization Problems and Risks

Despite the undeniable role of data visualization in providing an efficient tool for reasoning on quantitative data, there are a number of problems connected with the possible misuse of graphical data displays. Potentially, this can lead to a wrong interpretation of the message contained in the diagram.

In their books, Edward Tufte and Howard Wainer provided numerous examples of mistakes or intentionally introduced distortions in data visualization plots that can result in a message that is significantly different or even opposite to the one contained in the data. Most often these problems are related to misuse of axis scales, color palette, or elements of design.

Sabrina Bresciani and Martin Eppler provided a two-dimensional classification of data visualization problems. First, the problems can be induced by the designer (intentionally or unintentionally) or by the user of the diagram. Second, these problems can be classified into cognitive, emotional, and social ones. Cognitive problems can be connected to inappropriate use of graphical elements, lack of clarity, or oversimplification or overcomplexification of the graphical display or induced by heterogeneity of target user groups (e.g., women have a more accurate perception of the color palette than men). Emotional problems can be connected to a repulsive content of the graphical design. Social problems can be connected to cross-cultural differences of users (e.g., in some Eastern countries, time is shown from right to left, and the meaning of red and green is not identical to that accepted by Western countries).

In the field of data cartography, the main problem of data visualization is the possible distortion of mutual distances when projecting data points from a multidimensional to a low-dimensional space. Distant points in the multidimensional space can be projected at a short distance in the classical PCA method, and points close in the multidimensional space can be projected at a large distance in applications of nonlinear data–mapping techniques. There are visualization methods warning the user about possible distortions; however, in general, any conclusion derived from analyzing a data map should be verified carefully by rigorous statistical techniques of hypothesis testing.

Yet another source of problems in data visualization comes from the use of categorical or qualitative measurements for which no standardized and well-established graphical displays exist.

Data visualization risks should be distinguished from challenges posed to scientific data visualization in various fields. These challenges aim at making data visualization more informative and taking advantage of recent achievements in computer graphics, psychology, and computer science. For example, 10 top data visualization problems were formulated, such as the need for ameliorating usability and scalability of graphical displays as well as shifting the visualization focus from visualizing static structures to visualizing dynamics. Several international conferences (such as IEEE Visualization and IEEE Information Visualization) provide a wide forum for answering these challenges through international and interdisciplinary research efforts.

Andrei Zinovyev
Institut Curie
Paris, France

See also Correspondence Analysis; Data, Missing; Data, Spatial; Factor Analysis; Network Analysis; Time-Series Analysis

Further Readings

Bresciani, S., & Eppler, M. J. (2008). *The risks of visualization: A classification of disadvantages associated with graphic representation of information* (ICA Working Paper No. 1/2008). Lugano, Switzerland: University of Lugano, Institute for Corporate Communication, Faculty of Communication Sciences.

Brinton, W. (1939). *Graphic presentation.* New York: Brinton Associates. Retrieved October 13, 2010, from http://www.archive.org/details/graphicpresentat00brinrich

Gorban, A. N., Kegl, B., Wunch, D. C., & Zinovyev, A. (Eds.). (2008). Principal manifolds for data visualisation and dimension reduction. *Lecture Notes in Computational Science and Engineering, 58,* 96–130.

LeGates, R. (2005). *Think globally, act regionally: GIS and data visualization for social science and public policy research.* Redlands, CA: ESRI Press.

Lengler, R., & Eppler, M. (2007). Towards a periodic table of visualization methods for management. Paper presented at the *IASTED Conference on Graphics and*

Visualization in Engineering (GVE 2007), Clearwater, FL. Retrieved October 13, 2010, from http://www.visual-literacy.org

Melville, A., Polunin, Y., Ilyin, M., Mironyuk, M., Timofeev, I., Meleshkina, E., et al. (Eds.). (2010). *Political atlas of the modern world*. Malden, MA: Wiley-Blackwell.

Organisation for Economic Co-operation and Development. (2006). *OECD factbook 2006: Economic, environmental and social statistics*. Paris: Author.

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.

Wainer, H. (2009). *Picturing the uncertain world.* Princeton, NJ: Princeton University Press.

Ward, M. D. (2002). The development and application of spatial analysis for political methodology. *Political Geography, 21*(2), 155–158.

Websites

Data visualization portal: http://datavisualization.ch
The premier news and knowledge resource for data visualization and infographics.

Gapminder: Unveiling the beauty of statistics for a fact-based worldview: http://www.gapminder.org
An excellent dynamic and interactive data visualization tool and a database of socio-economical indicators collected for 200 years of world history.

Google Public Data Explorer: http://www.google.com/publicdata
Makes large data sets easy to explore, visualize, and communicate. The public data sets available contain several fact books (such as OECD Factbook*) that can be used in political science studies.*

Human Development Reports: http://hdr.undp.org/en/reports
Contains tons of examples of thoughtful data visualization, especially in recent reports.

Mapping Globalization Project of Princeton University: http://qed.princeton.edu/main/MG/Data_and_Analysis
The aim of the project is to visualize trading relations between countries.

DECENTRALIZATION

The concept of decentralization refers to the allocation of power in organizations or social structures usually from the higher to the lower-level structure(s)/organization(s). It can describe either an existing structure in which smaller or peripheral units have effective powers or a process of structural change implying a shift of power from the center to these units. The structure concerned can be a network differentiated into a center and a periphery, a hierarchical organization internally differentiated into sectors or subdivisions, or a territorial organization differentiated into levels of geographical space.

In political science, decentralization usually refers to multilevel structures of government or administration. It results from (re)allocation of power to elect or denominate policymakers of legislative or administrative competences or of fiscal resources from higher to lower levels. Given interdependencies between levels, decentralization is to be regarded as a relational concept. The effects of policies made by lower-level units depend on the power of the center, and shifts of power toward lower levels usually affect the degree of centralization or decentralization, while only major reforms change the character of a political system, for example, by turning a unitary into a federal system. Nevertheless, regardless of its extent, decentralization affects governance and democratic legitimacy. For this reason, it is not only a matter of multilevel power games but also contested for normative reasons.

For long, decentralization was not a salient concept in political theory. History of government was about centralization of power to manage conflicts of competing local elites and social groups. The modern state evolved in a process of center formation. The 20th century saw a trend toward inter- or transnational governance. However, these developments never went in only one direction. Centralization was also often thwarted by countervailing powers from below. In the final decades of the 20th century, a general trend toward more decentralization gained ground in many developed states. At the same time, decentralization became a kind of paradigm for governance in developing countries.

In political theory and practice, decentralization is now mostly regarded as a preferable alternative to centralization. As a principle guiding the organization of government, it is justified by different normative theories. All of them can be traced back to different political theories and ideas, but all are under dispute, as is exemplified by the writings of Johannes Althusius and Jean

Bodin or the controversies between the American Federalists and anti-Federalists.

The first reason for decentralization is derived from the principle of subsidiarity. Rooted in the political philosophy of reformed Protestantism in the early 17th century and in social theory of the Catholic Church of the 19th century, it defends the priority of small social communities against large societies and a centralized, bureaucratic state. While this collectivist view reemerged in communitarist theories, the principle has a strong impact in discourses on constitutional design. Here, subsidiarity has turned into a legal principle stipulating a prerogative for decentralization in case of dispute.

To justify this prerogative, two main arguments are relevant: The first holds that decentralization improves democracy. If government is closer to the people, communication between citizens and their representatives should be more effective and individual preferences should have better chances to be considered in political decisions. Moreover, decentralization is said to reflect plurality and regional differentiation of societal interests. Liberalist theories prefer decentralization as a device against dictatorship either of autocrats or of majorities. Against these assumptions, empirical studies have revealed, that governance at lower levels is often more elitist and exclusive compared with representative democracy in larger territorial units and that it may, under particular conditions, support regional conflicts or separatism.

Economic theories stipulate that decentralization fosters efficiency of policy making. If citizens have an opportunity to choose between alternative supplies of public goods or services provided by lower-level governments demanding different tax burdens, they force governments competing for taxpayers to make every effort to find an optimal ratio of costs and benefits. However, to achieve this positive effect, two basic conditions must apply: First, a sufficient number of taxpayers from different groups of a society must be able and willing to move between jurisdictions. Second, policies should not produce external effects spilling over the borders of a jurisdiction. In practice, moving to change jurisdictions ("voting by feet") is hardly a real opportunity for ordinary citizens, and external effects occur frequently. Therefore, to encourage innovation, economic theory now emphasizes "yardstick competition," in which suppliers of goods and services are evaluated against best practices in that economic sector. But again, this outcome can be achieved only under particular provisions. Moreover, if policies have redistributive effects, decentralization is problematic. In regionally divided societies, it can cause vicious cycles leading to inequality, inefficiency, and instability.

Finally, political scientists praise decentralization as a device to avoid a concentration of power in a government. Powers should be divided not only between executive, legislative, and judiciary but also between central and regional or local levels. Although the general argument is convincing, it leaves open the question of which powers should be centralized and which should be decentralized. Moreover, decentralization can divide powers for legislation, as is typical in a federal state, or it can refer to executive powers while legislation remains to the center, as is often the case in unitary states. There is no convincing evidence to conclude which system creates better checks and balances and which one is protected better against centralist trends.

Normative disputes on the effects of decentralization can be settled by empirical research. However, decentralization is difficult to measure. First, as mentioned above, power is a multidimensional concept. It can be based on formal competences, on effective capacities to make or impede decisions, on support by electors/voters or influential associations, or on fiscal resources. Usually, scholars find no congruence in the allocation of these different sources of power, and weighting their relative significance is rather problematic. Second, the degree of centralization and decentralization can vary between the legislative, executive, and judiciary branches of government. The party system reflecting the relations between government and society may also matter when measuring decentralization and may differ from other institutional dimension of multilevel government. Third, to measure the degree of decentralization, researchers have to assess the power relations between levels of government. Decentralization is often identified with "self-rule,"—that is, the power of regional or local governments to decide on their affairs autonomously. But usually, central governments can intervene in lower-level policy making by different means such as regulation, supervision, incentives, or capacity building. On the other hand, decentralized authorities may be able to evade central control.

Also, in federal systems, regional governments (i.e., representatives from states, *Länder,* or provinces) participate in decisions of the central government ("shared rule"). Finally, the real power generated by decentralization of formal competences or resources depends on governance capacities of lower levels. Under certain condition, decentralization may overburden governments with unsolvable conflicts and problems and at the same time unburden the central government from pressure.

Nevertheless, by constructing indicators, scholars have been able to provide information about decentralization in governments. In contrast to economics, where fiscal issues are considered, political scientists try to measure relative political power of governments. By applying their Regional Authority Index, Gary Marks, Liesbeth Hooghe, and Arjan H. Schakel (2008) recently revealed that, among the 42 advanced industrialized countries covered in their study, 29 saw an increase of regional powers, while only two countries saw a decline. Studies on local government have portrayed a similar evolution, although often with the qualification that many cities have suffered from fiscal constraints during the last decades. In general, decentralization seems to constitute a general trend in government, at least in developed countries.

This trend can be explained by different factors. One is the change in public policies, in particular, the shift from distributive welfare policies to the provision of services. The process of democratization is also said to have supported decentralization. Many governments have to respond to pressure from new nationalist or regionalist movements demanding political autonomy. Economic globalization has reduced the effects of national policies to support firms and increased the relevance of regional and local responsibilities for infrastructure and administrative services. Discourses about public sector reform have been guided by the paradigm of decentralization, which influenced institutional reforms in many countries. Besides these forces toward decentralization, institutionalist theories rightly point out that the "stickiness," the ability to succeed, of existing institutions constrains a reallocation of power. In multilevel governments, proposals to decentralize raise conflicts between central, regional, and local governments. Even if external forces or approved reform concepts support decentralization, central governments try to maintain their powers. Therefore, the outcome of processes of decentralization usually reflects compromises between actors involved in bargaining for power, which is often perceived as a zero-sum game. However, these explanations do not add up to a theory and most of the individual hypotheses are still not sufficiently tested in empirical research.

Unlike what is suggested in normative reasoning, there is also no theory explaining the effects of decentralization, whichever aspect is under consideration. On the contrary, as Daniel Treisman has shown in a comprehensive review of the relevant literature, available empirical studies provide at best mixed results. What decentralization brings about depends not only on the particular mode of decentralization but also on additional conditions of governance, such as the patterns of democracy, organization of society, economic situation, or political culture. As a rule, benefits of decentralizations have to be charged up against the costs, and the evaluation of the balance depends on interests, standards, and perspectives. For this reason, decentralization should not be taken as a value or principle for organizing government. Rather, it should be regarded as an issue in institutional or constitutional policy, which, from a normative point of view, requires a careful study of positive and negative effects under particular conditions, and, from an analytical point of view, is a matter of political conflicts and decisions making.

Arthur Benz
Technische Universität Darmstadt
Darmstadt, Germany

See also Autonomy, Subnational; Balance of Power; Federalism; Governance, Multilevel; Governance, Urban; Institutional Theory; Local Government; Regionalization

Further Readings

Cheema, G. S., & Rondinelli, D. A. (Eds.). (2007). *Decentralizing governance: Emerging concepts and practices*. Washington, DC: Brookings Institution Press.

Marks, G., Hooghe, L., & Schakel, A. H. (2008). Patterns of regional authority. *Regional and Federal Studies, 18,* 167–181.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of Economic Literature, 39,* 1120–1149.
Rodden, J. A. (2004). Comparative federalism and decentralization: On meaning and measurement. *Comparative Politics, 37,* 481–499.
Treisman, D. (2006). *The architecture of government: Rethinking political decentralization*. Cambridge, UK: Cambridge University Press.

Delegation

Delegation occurs in politics whenever one actor or body grants authority to another to act on behalf of or to carry out a function for the first in a political process. In such general terms, delegation is ubiquitous and a defining feature of politics beyond direct individual actions. Voters delegate to elected officials in representative government; governments delegate to ambassadors in foreign affairs; legislatures delegate to committees the authority to study policy issues and report bills and to bureaucracies the authority to make policy.

Because of the breadth of the topic, this entry focuses specifically on delegation from legislatures to bureaucracies in administrative and bureaucratic governments. Delegation has become inherent in this mode of governance as the reach of public policy has expanded beyond what elected legislatures can possibly handle. Such delegation presents particularly interesting institutional and political problems in the United States due to separation of powers, and thus, this entry focuses even more specifically on the rationale for and determinants of delegation in this case. Because of space limitations, this entry does not address whether agency use of delegated authority is responsive to the policy goals of external political actors (Congress, the president, courts, interest groups) or the channels by which this responsiveness is effected.

Scope of Delegation to Bureaucracies in the United States

Delegation to bureaucracy presents important questions of political and democratic legitimacy because it often imparts to the bureaucracy some power to shape the law. In practice, delegation involves the bureaucracy in making law, beyond the nonlegislative functions of executing laws passed by Congress. Yet Article 1 of the Constitution explicitly allocates the power to make law only to Congress. Moreover, delegation of lawmaking power implies that bureaucracies blend the formal powers (legislative, executive, and judicial) that the Constitution separates. The Supreme Court held such delegation, in its most expansive form, unconstitutional in its 1935 decision in *Schechter Poultry Corporation v. United States*, which invalidated the administrative nerve center of Franklin Delano Roosevelt's New Deal. Since then, the typical formulation to legitimize Congress's delegation of its Article I power to make law is that Congress articulates at least a "general standard" to guide agency policy making, and agencies merely "fill in the details." However, no statute has been invalidated by the Supreme Court on delegation grounds since 1935; even extremely broad guidelines stipulating only that agencies regulate "in the public interest" have been upheld. Thus, it is debatable whether Congress faces any meaningful legal restrictions on its delegation to agencies.

James M. Landis, one of the foremost legal scholars of regulation and architect of the Securities and Exchange Commission, disputed such critiques of delegation on the grounds that, first, each constitutional branch of government possessed ample checks over administrative agencies, and second, delegation to administrative boards was necessary to reconcile democratic government with the formidably complex policy problems created by economic developments. This defense has retained intellectual currency since Landis first offered it.

Causes and Effects of Delegation

The history of delegation to bureaucracy in the United States is a history of Congress working hard to give away a measure of formal power allocated to it in the Constitution. Since Congress often keeps close watch on its constitutional prerogatives, this is a choice that is interesting to try to explain. Scholars have offered a wide variety of theories to do so. Broadly speaking, the contemporary political science literature avers that delegation of policy-making authority to bureaucracies allows a legislature to achieve policy ends that it could not achieve through legislation itself.

First, Congress may delegate simply because of the opportunity cost of its time spent on any one issue. For example, in policy pertaining to the electromagnetic spectrum (EMS), Congress could set technological standards and allocate spectrum rights to various classes of users itself. Instead it has charged the Federal Communications Commission (FCC) to do so "in the public interest." Congress can therefore reserve its attention for broader and more nationally significant policy choices, leaving the FCC with the relatively more arcane choices over EMS use.

This rationale suggests that delegation presents Congress with a "principal–agent problem." The bureaucratic agent may pursue different goals with its authority from those Congress (the "principal") would pursue. Several scholars have noted that legislative and executive policy goals are more congruent under unified than under a divided government. This reasoning has led to one of the more robust empirical findings on delegation: legislative delegation of policy-making authority to bureaucracies increases under a unified government. David Epstein and Sharyn O'Halloran (1999) established this at the federal level in a study of delegation from 1947 to 1995. John Huber and Charles Shipan (2002) further established that this pattern holds at the state level in the United States and comparatively among the Organisation for Economic Co-operation and Development countries (where "policy conflict" between the legislature and executive is operationalized by minority coalition governments).

A second and closely related explanation for delegation is that Congress delegates to tap into expertise that bureaucrats have over some policy domain. For instance, both the FCC and the Department of Justice make policy with respect to market competition; by delegating this authority, Congress can leverage the expertise these organizations have developed in the economic analysis of the competitive effects of industrial structure. The implicit presumption is that Congress does not have or wish to develop this expertise itself, which at some level is presumably a result of the opportunity cost of time spent doing so.

Formally, delegation of policy-making authority is often conceived as granting a "zone of discretion" to an agent, a set of policies among which the agent can choose freely and without further interference from the legislature. The legislature decides which policies are in the zone and which are not. This was formulated in 1984 by Bengt Holmström and was first applied in the political science literature on bureaucracy by Epstein and O'Halloran (1994). The critical strategic problem confronted by Congress is to induce the agency to use its expertise as Congress itself would use it, if Congress possessed it, by tailoring the zone of discretion in particular cases. Because the zone of discretion is a blunt instrument, this problem generally cannot be perfectly "solved" from Congress's point of view; it is inherent that the agent will sometimes use its information to pursue policies Congress does not always prefer. This "agency loss" is simply a necessary cost of an otherwise desirable process and does not undermine the case for delegation in normative terms. A Congress aware of this agency loss would simply not delegate if it were too large compared with the benefits of delegation.

Empirical evidence bears out that bureaucratic expertise is a rationale for delegation. For instance, in their comprehensive study of major postwar federal legislation, Epstein and O'Halloran (1999) found that laws pertaining to more technically complex policy areas delegated more authority to bureaucratic agencies, net of constraints simultaneously placed on those agencies, than laws pertaining to less complex policy areas.

Of course, agencies cannot always be presumed to possess expertise, especially when they are initially created to deal with a novel policy problem. Delegation can also help induce agencies to develop expertise. First, agencies have an incentive to develop expertise if they anticipate that their policy authority over tasks they deem important will increase as a result (Sean Gailmard & John Patty, 2007). Second, delegating authority over a task that an agent considers important provides an incentive for the agent to acquire information about how best to perform the task. Given the delegated authority over an important task, the agent remains ignorant at its peril; acquiring expertise allows the agent to obtain private benefits from good performance.

Third, delegation can act as a commitment by Congress to a future course of policy. For instance, under the Steel Trigger Price Mechanism, the

International Trade Commission (ITC) investigates possible instances of "dumping" in the United States by international trading partners. The policy commits the United States to import restrictions whenever dumping is found. If Congress had not delegated this authority to the ITC but instead retained the authority to determine policy on a case-by-case basis, import restrictions in response to dumping would not be so evident. It would depend on the policy preferences of members of Congress and the president toward import restrictions. A protrade Congress or president would be sufficient to prevent enactment of import restrictions, even if illegal dumping were identified. By delegating fact-finding authority and programming, a Congress at one time can preprogram the policy response to changing conditions even after a new Congress is seated. In celebrated analyses of the political origins of administrative institutions, Matthew McCubbins, Roger Noll, and Barry Weingast (1987) offer separate but related arguments that commitment of future policy is an important rationale for delegation. Such delegation as commitment is only effective if the administering agency is able to hold steady in changing political winds; thus, T. M. Moe predicts that in such cases agencies will be deliberately insulated from and unresponsive to future Congresses and presidents.

Opportunity cost theories suggest implicitly that delegation inherits whatever normative appeal is inherent in congressional policy goals. Commitment theory is more ambiguous because it suggests that one legislative coalition achieves its goals by compromising the ability of future coalitions to do the same, thus dampening the responsiveness of public policy to shifts in congressional goals.

Besides explanations based on achieving policy ends, scholars have also argued that delegation allows legislators to evade accountability and difficult policy choices. Congress may delegate to shift blame if policy may be unpopular or fails to achieve desired results. In some policy domains, success is expected and failure is a high-profile event. In airline safety, the public does not celebrate flights that land safely. When airplanes operate safely the public may even protest that safety and security measures are too onerous in terms of time and money. But when airplanes crash we can expect recriminations about gaps in the system. Much the same story applies to homeland security issues. The blame-shifting explanation suggests that Congress delegates authority over these policy areas so that Congress will not be held responsible in case of a problem, and doing so may even allow Congress to claim credit for solving problems. This explanation implicitly assumes the public does not realize that Congress chose to arrange the system this way and is thus still ultimately responsible. For instance, if the decision to delegate is separated by a long interval from resulting problems, it may not make sense for voters to hold sitting legislators answerable for the delegation and oversight of the agency's initial policy choices.

Conclusion

Delegation from Congress to bureaucratic agencies, including the delegated power to make law, is pervasive in the United States. Indeed, it is a given in most major legislation passed since the New Deal. Delegation to administrative agencies cannot remove politics from policy making; it can merely move the political choices around in the policy process and push politics inexorably further into administration. Delegation is a choice that requires both normative justification and positive explanation. The positive explanations reviewed above are not necessarily mutually exclusive, though they carry dramatically different normative implications for how delegation from legislatures to bureaucracies should be evaluated.

Sean Gailmard
University of California, Berkeley
Berkeley, California, United States

See also Policy Process, Models of; Regulation

Further Readings

Epstein, D., & O'Halloran, S. (1994). Administrative procedures, information, and agency discretion. *American Journal of Political Science, 38,* 697–722.

Epstein, D., & O'Halloran, S. (1999). *Delegating powers: A transaction cost politics approach to policymaking under separate powers.* New York: Cambridge University Press.

Gailmard, S., & Patty, J. W. (2007). Slackers and zealots: Civil service, policy discretion, and bureaucratic expertise. *American Journal of Political Science, 51*(4), 873–889.

Holmström, B. (1984). On the theory of delegation. In M. Boyer & R. Kihlstrom (Eds.), *Bayesian models in economic theory.* Amsterdam: Elsevier Science.
Huber, J., & Shipan, C. (2002). *Deliberate discretion: The institutional foundations of bureaucratic autonomy.* New York: Cambridge University Press.
McCubbins, M., Noll, R., & Weingast, B. (1987). Administrative procedures as instruments of political control. *Journal of Law, Economics, and Organization, 3*(2), 243–277.
Schechter Poultry Corporation v. United States, 295 U.S. 495 (1935).

Deliberative Policy Making

Deliberative policy making applies principles derived from the theory of deliberative democracy to public policy making. The theory of deliberative democracy responds to the perceived shortcomings of a majoritarian democracy where actors bargain to defend their interests (strategic action), bargaining and voting procedures do not change actors' preferences while discouraging social learning, the strong impose their will on the weak, and constraints on lying, deception, and manipulation are few and far between. Against this vision of democracy as a pluralist interest aggregation, deliberative democracy upholds an alternative model based on discussion and persuasion where actors must defend and criticize proposals with reasoning they believe others will accept (communicative action), public discussion can transform actors' preferences, the majority prefers to compromise with the minority, and inclusion in public discussion of all social positions and perspectives aims to maximize social learning. Deliberative policy making aims for genuine preference transformation rather than mere preference aggregation by expanding in governance institutions the role of deliberation based on norms such as reason giving, publicity, joint problem solving, and inclusive participation. This entry reviews the aims and varieties of deliberative policy making before assessing its achievements and shortcomings.

Aims

Deliberative policy making seeks to reach policy decisions that are both democratically legitimate and technically sound. Especially as applied to environmental regulation, theorists such as Walter Baber and Robert Bartlett have defended the substantive claim that expanding deliberation in policy making increases the ecological rationality of decisions. Unlike traditional, hierarchical forms of administrative decision making, deliberative, horizontal methods are said to be more flexible and robust in the face of the epistemic complexity and uncertainty that characterizes such policy problems in the "network society." By revitalizing the role of nonexpert citizens in policy inquiry and generating policy-relevant information on as wide a social basis as possible, deliberative policy making helps guard against closing off debate prematurely and discarding policy alternatives.

A normatively demanding conception of public deliberation, which theorists argue is superior to pluralist bargaining, is said to promote these aims. Public deliberation as conceived by theorists such as James Bohman is not merely talk; rather, it involves the exchange of reasons aimed at evaluating alternative proposals for action. Citizens submit their ideas and beliefs for discussion and criticism by fellow citizens, making it less likely that selfish or poorly thought-out proposals will survive the debate. Note that public deliberation requires a considerable level of freedom and equality: inclusion of everyone affected by a decision, substantial political equality including equal opportunities to participate in deliberation, equality in methods of decision making and setting the agenda, free and open exchange of information and reasons sufficient to acquire an understanding of both the issue in question and the opinion of others. These conditions are much more restrictive than the standards usually applied to political debate and difficult to achieve in practice.

Efforts to adapt such principles of public deliberation for practical use have yielded public consultation techniques that go beyond government agencies merely seeking public input from individual citizens or firms. Rather, their goal is to generate high-quality deliberation conceived as dialogue among the participants. These techniques include the consensus conferences pioneered in Denmark, the deliberative opinion polls developed by James Fishkin, collaborative governance schemes involving multiple stakeholders, and the citizen jury. The next section classifies these varieties of deliberative policy making, using a simple typology.

Varieties: Hot Versus Cold, Weak Versus Strong

Deliberative policy-making techniques vary along two dimensions: partisanship and authority. Deliberative forums differ widely in how they confront the question of partisanship. Some deliberative forums seek to accommodate partisanship by inviting partisan stakeholders to participate in the deliberations and then structuring the proceedings so as to maximize deliberation while minimizing bargaining (consensus conference, collaborative governance). Stakeholders are included in such forums on the assumption that decisions reached without their participation lack the broad-based acceptance required to translate decisions into practice. Because they aim to accommodate pluralism, partisan forums may be thought of as a deliberative *extension* of pluralist interest aggregation. Nonpartisan forums, by contrast, seek to minimize partisanship by excluding interest groups from the deliberations (deliberative opinion poll, citizen jury) in the expectation that interest representation inevitably diminishes partisans' ability to follow the norms of deliberation outlined above. Here the goal is to approximate Jürgen Habermas's "ideal speech situation," which brackets forms of manipulation and coercion external to the logic of reasoned argument and persuasion. Insofar as they hope to transcend pluralism, nonpartisan forums may be thought of as a deliberative *alternative* to pluralist interest aggregation. The respective virtues of "hot" and "cold" forms of public deliberation are subject to increasingly refined empirical analysis in the literature.

A further variation among forms of deliberative policy making is their degree of political authority. This is expressed as the difference between "strong" and "weak" publics in the literature. Weak publics such as the informal public sphere engage only in opinion formation, while strong publics such as parliaments also make collectively binding decisions about the proposals before them. Endowing nonelected, nonexpert, ad hoc deliberative forums with the authority to make policy decisions is a contentious proposition, however. Is it a permissible delegation of legislative authority, or does it violate basic norms of representative democracy? In practice, weak publics predominate. Deliberative policy making is typically advisory in nature; participants are confined by sponsoring agencies to preparing recommendations as a basis for decision making by elected officials and public managers.

Deliberative forums can be evaluated in terms of their deliberative capacity, impact, and legitimacy. High-quality deliberation may be more easily achieved in nonpartisan forums, while partisan forums may have a greater impact on policy decisions and enjoy greater legitimacy—especially where external conditions provide sufficient incentives for powerful stakeholders to prefer deliberation to other ways of resolving their differences, such as litigation. However, given the tremendous variability, first, among designed deliberative forums and, second, among the political, administrative, and legal contexts in which they must function, these empirical relationships are not particularly robust. The complex interplay between deliberative designs, their internal dynamics, and external context limits the potential for generalizing about deliberative policy making.

Critiques of Deliberative Policy Making

This section considers some popular objections to deliberative policy making—namely, that its conception of democratic legitimacy is substantively too demanding, that it is too costly, and that it cannot accommodate interests.

Deliberative Policy Making Asks Too Much of Democratic Legitimacy

It is rarely realistic, critics contend, to expect reasoned debate to lead to a consensus supported by all participants. Public deliberation, however, does not demand that all participants support an agreement for the same reasons, or that they come to any agreement at all. Bohman's "plural agreement," for example, only requires a commitment to continued cooperation in public deliberation, even with persistent disagreements. The very conditions of public deliberation—publicity, reversibility of decisions, and social inclusiveness—encourage an ongoing cooperation with others who disagree, which is at least not unreasonable. The give-and-take of reasons in public deliberation makes it less likely that irrational and untenable arguments will decide outcomes. Decision-making procedures allow revisions of arguments and decisions and even include procedures that either take up features of defeated positions or improve their chances of being heard. And finally, deliberative decision-making procedures are broadly inclusive, so minorities may hope to affect future outcomes.

The condition of publicity ensures that decisions are more likely to be made on the basis of all relevant perspectives, interests, and information and that they are less likely to exclude legitimate interests, relevant knowledge, or dissenting opinions. When proposals are subjected to a wide range of possible alternative opinions in a forum that is open to all societal interests and perspectives, the quality of reasons for decisions is therefore likely to improve. Participants—and, it is hoped, citizens generally—should be able to accept decisions reached in this manner as reasonable unless and until they are shown to be otherwise in future rounds of deliberation.

Deliberative Policy Making Is Inefficient

Deliberative policy making requires additional resources, increasing administrative costs. Insofar as decisions reached through deliberation can be expected to enjoy broader public support than decisions reached by nondeliberative means, they may avoid costly and time-consuming legal challenges and delays in implementation. At the very least, deliberative policy making provides administrators with better information. Adding the cost of deliberation to an already complex decision-making process may improve the odds that a given decision will survive legal scrutiny. In the United States, for example, most Environmental Protection Agency regulations face court challenges at some point in their history. Even major regulations, often the result of several years of work, do not always survive legal scrutiny. More and better public deliberation in rule making makes such failures less likely because it improves the odds that all relevant perspectives, interests, and relevant objections will be considered from the outset.

Deliberative Policy Making Cannot Accommodate Interests

Is deliberative policy making naive about the role of interests, and hence of negotiation and bargaining, in politics? Deliberative democracy makes two claims about interests, neither of which denies its role in democratic politics. The first is that interests cannot be known prior to public deliberation. On this view, interests are not fixed or given but subject to discovery and transformation through public debate about joint problems and conflicts. In the absence of public deliberation, therefore, interests cannot be adequately taken into account in decision making. The second claim is that policy proposals cannot be defended merely on the basis of self-interest because such a defense is unlikely to survive deliberative scrutiny. Participants will be suspicious of proposals that serve selfish purposes but are falsely advanced in the name of public interests. There is no denying in deliberative policy making that politics involves interests; rather, it encourages their disclosure and sometimes their transformation.

Christian Hunold
Drexel University
Philadelphia, Pennsylvania, United States

See also Democracy, Theories of; Discursive Institutionalism

Further Readings

Baber, W. F., & Bartlett, R. V. (2005). *Deliberative environmental politics: Democracy and ecological rationality*. Cambridge: MIT Press.

Bohman, J. (2000). *Public deliberation: Pluralism, complexity, and democracy*. Cambridge: MIT Press.

Fiorino, D. (2006). *The new environmental regulation*. Cambridge: MIT Press.

Fischer, F. (2009). *Democracy and expertise: Reorienting policy inquiry*. Oxford, UK: Oxford University Press.

Democracy, Consociational

Consociational (from Latin, *consociatio*—close political and social interlinkage) democracy denotes a democratic political order in which political decisions are reached mainly by negotiations between political elites. Majority rule is not the dominant technique in decision making. Rather, political actors seek to find broad compromise based on amicable arrangement. Switzerland and the Netherlands are frequently cited as examples of consociational democracy. There is a variety of conceptual definitions of consociational democracy. However, consensus exists about the antipode. This is a majoritarian democracy in which elites eschew negotiations and compromise as soon they have a plurality of 50% of the vote plus one. The best known empirical illustration is the political system of the United Kingdom. Therefore, the antipode to consociational democracy is frequently labeled "Westminster democracy" since the British parliament resides in the Palace of Westminster. The political system of postwar New Zealand comes even closer to the ideal type of a majoritarian democracy.

The concept of consociational democracy is important for the comparative analysis of political systems since it emphasizes variations between forms of democratic government. This is in stark contrast to the traditional argument that English-speaking countries have a homogeneous political culture—that is, the actors share general orientations in the political system. There is consensus with regard to general political ends and means. This is a prerequisite of a stable democratic order. The more homogeneous a political culture is, the higher the stability of democracy, with the Anglo-Saxon countries being particularly homogeneous. In a nutshell, there is one form of democratic government, and the Anglo-Saxon countries are the best examples. In contrast, Continental European countries have a heterogeneous political culture. Subgroups such as Catholic groups, workers' movements, and the middle classes differ considerably in ideology and strategy. These fundamental differences are the germ of system instability.

However, postwar history demonstrated that, in deeply divided European societies, heterogeneity in political culture and a stable democracy are not mutually exclusive. The concept and underlying causal hypotheses of consociational democracy provided explanations for the political stability in segmented societies. Obviously, there are different routes to the Rome of democracy and consociational democracy is one of these variants.

Varieties of Consociational Democracies

The term *consociational democracy* entered the scientific debate in 1961, when it was used to describe African political systems that accommodated by compromises a large number of groups with divergent ideologies to achieve a common goal. Arend Lijphart and Gerhard Lehmbruch applied the term to Western European countries at the end of the 1960s. In his book about the Dutch political system, Lijphart argued that, in a critical historical period during World War I, the leaders of the Dutch societal segments agreed on pragmatic solutions to deep conflicts over the relation of state and churches in education and over the electoral formula. This *pacificatie* (peaceful settlement) of 1917 had three major characteristics: (1) the elites negotiated on behalf of their segments without any extensive consultation with their constituencies, (2) elites of all segments participated in the agreement, and (3) the principle of proportionality in the substance of the settlement was of utmost importance (Lijphart, 1968). A deliberate decision of the elites to cooperate as the best response strategy underlay this *pacificatie*. Once practiced, this mode of top-level compromise became institutionalized in the Netherlands. The basic sociocultural prerequisite of this system was the pillarization (*verzuiling*) of society. Individuals belong to clearly separated societal pillars—that is, sociocultural subgroups—and only the elites on the top of these pillars had contact with and negotiated with elites from other pillars. Operational criteria for the classification as a consociational democracy are grand coalitions, cultural autonomy of subgroups, proportionality, and minority voting (Lijphart, 2008).

Lehmbruch (1967) analyzed the pattern of conflict management in Austria and Switzerland. In Switzerland, the important conflicts were those that emerged between groups that are regionally concentrated: cantons with Catholic or Protestant orientation, agricultural or industrial structures, or German- or French-speaking populations. Hence bargaining between regional elites was the dominant pattern of conflict regulation. In contrast, Austria resembled the Netherlands in the sense that the conflicts existed between sociocultural groups—Catholic-Conservative versus Socialist camps—without a regional identification. Lehmbruch used the Austrian term of *Proporzdemokratie* (democracy based on proportionality) for consociational democracy; he also pointed to the Swiss term of *Konkordanzdemokratie* (democracy based on concordance). Hence, consociational democracy, Konkordanzdemokratie, or Proporzdemokratie can be used synonymously to describe a form of government in which the majority rule is not the dominant technique of decision making. The underlying reason is a deep sociocultural or political division of the country that does not allow for a system of alternating majorities. Either there is no group that could be a majority or the minority groups are very suspicious of the other groups and see little chance of becoming a majority sometime in the future.

Obviously, consociational democracy is hard to define in operational terms. This may be one reason for the development of a new concept—consensus democracy—which is described as a system of power sharing. Lijphart (2008) points to four major differences between consociational and consensus democracy. The first difference is that *consociational democracy* denotes the peaceful system of conflict regulation in deeply divided societies. The term condenses the main findings from the descriptive analysis of these political systems. In contrast, *consensus democracy* started from the ideal type of a majoritarian democracy, enumerating all the major characteristics and then defining the nonmajoritarian as a contrast. This led to the following list of variables. It can be subdivided into the two dimensions of "Parties–Executives" and "Federal–Unitary." These dimensions are, however, only weakly correlated empirically (Lijphart, 1999).

Parties–Executives

1. Concentration of executive power in single-party majority cabinets versus executive power-sharing in broad multiparty coalitions.
2. Executive-legislative relationships in which the executive is dominant versus executive–legislative balance of power.
3. Two-party versus multiparty systems.
4. Majoritarian and disproportional electoral systems versus proportional representation.
5. Pluralist interest group systems with free-for-all competition among groups versus coordinated and "corporatist" interest group systems aimed at compromise and concertation. (pp. 3–4)

Federal–Unitary

1. Unitary and centralized government versus federal and decentralized government.
2. Concentration of legislative power in unicameral legislature versus division of legislative power between two equally strong but differently constituted houses.
3. Flexible constitutions that can be amended by simple majority versus rigid constitutions that can be changed only by extraordinary majorities.
4. Systems in which legislatures have the final word on the constitutionality of their own legislation versus systems in which laws are subject to judicial review of their constitutionality by supreme or constitutional courts.
5. Central banks that are dependent on the executive versus independent central banks. (pp. 3–4)

The second difference concerns the principles of consociationalism, which are broader than those of consensus democracy. For instance, in the consociational framework, *grand coalition* could mean more than just a multiparty coalition; grand coalition could be created informally by advisory arrangements or by alternating presidencies. This is connected to the third difference: Consensus democracy tends to emphasize formal institutional devices while consociational democracy relates mainly to informal practices. Finally, consociational democracy is more demanding in the sense that group autonomy or inclusion of all relevant groups is required while consensus democracy provides incentives for cooperative behavior.

To avoid the normative undertones and evaluations of the term *consensus democracy,* André Kaiser (1997) suggests the term *negotiation democracy*. This seems to be particularly appropriate since *consensus democracy* combines basically three different aspects of coordination: between political elites (which is the sole focus of consociational democracy), between governments and interest groups (which is the focus of the debate on corporatism versus pluralist interest representation), and between veto points that may have competitive or collective orientations. According to Vicki Birchfeld and Marcus Crepaz (1998), competitive veto points occur when different political actors operate through separate institutions with mutual veto powers, as in federalism and strong bicameralism. Collective veto points emerge from institutions where the different political actors operate in the same body and whose members interact with each other on a face-to-face basis, as in proportional electoral systems and multiparty governments.

Genesis and Determinants

The literature distinguishes between at least three causal explanations of consociational democracy. The institutional explanation points to incentives and constraints produced by institutions, such as electoral laws or direct democracy. Electoral rules allowing for proportional representation tend to lead to multiparty systems. The need to negotiate and to build coalitions is larger in such party systems since there is no clear majority party that is able to dominate all the other parties. Likewise, direct democracy creates strong incentives for elites to cooperate and to build compromises. If they try to push through their goals, they risk the possibility that neglected groups may trigger a popular vote and that the policy of the majority group in parliament or government may be terminated by a popular vote. This is one reason why Switzerland embarked on conflict regulation by negotiation and inclusion of all groups that are capable of triggering a popular vote.

A second explanation relates to deliberate decisions of the elites (Lijphart, 1968). According to this explanation, in a critical historical situation, elites consider their options. They then take a deliberate decision for compromise solutions—and waive the option of conflictual strategies—based on either careful cost–benefit analyses or orientation to the common good. Once the basic decision for negotiation is taken, it becomes institutionalized. Lijphart (2008, pp. 51–52) identified nine facilitating conditions that make this decision more likely:

1. the absence of a solid majority that prefers pure majority rule to consociationalism;
2. the absence of large socioeconomic differences between the subgroups of the society;
3. a moderate number of groups, so that negotiation is a feasible option and is not made too difficult and too complex;
4. the groups having roughly the same size;
5. a relatively small population—such as in the classical cases of European consociational democracy (Austria, Belgium, Luxembourg, the Netherlands, and Switzerland);
6. an external economic or military threat that promotes internal unity;
7. overarching loyalties, such as a common national identity;
8. federalism as a means to foster group autonomy (if groups are regionally concentrated), which is conducive to consociationalism; and
9. traditions of compromise and accommodation.

Controlling for these favorable conditions, the core argument of this explanation is a voluntaristic, deliberate choice by elites. Hence, in principle, the option of consociationalism exists in all divided societies; it just depends on elites who are convinced of this type of conflict regulation.

A third explanation emphasizes the role of history and learning. Consociational democracy is more likely when elites are used to compromise and negotiate. This tradition amplifies learning processes. Elites become socialized in systems of negotiations, and they learn that negotiation and accommodation are appropriate techniques of decision making. Alternative conflictual decision modes are normatively inferior, less efficient, and more costly in several respects. Lehmbruch has radicalized this explanation by arguing that the consociational and corporatist European systems can be traced back to the religious peace treaties of the 17th and 18th centuries.

Impacts and Consequences of Consociational Democracy

Consociational and consensus democracy have a mixed balance of impacts and consequences. Manfred Schmidt (2008, pp. 313–316) identified six deficits:

1. Consociational techniques depend on top-level negotiations and compromise between elites.

Most of this accommodation is done behind closed doors; otherwise, politicians will not have enough flexibility given the constraints of electoral competition. By implication, voters cannot check whether their representatives stay within the corridors of policy choices that conform to the preferences of the electorate. Since the compromises are considered to be a result of quid pro quo, it is hard to assign responsibility to particular political groups and parties.

2. These compromises require that elites can decide on behalf of their constituencies and that these constituencies follow passively the leaders of their sociocultural group. Politically interested and active individual citizens are a disturbance variable in consociational democracies.
3. Minority groups have the right to veto and can obstruct solutions that are in the interest of the overwhelming majority of the nation. While competitive democracy has to face the danger of the tyranny of the majority, consociational and consensus democracies may suffer from the tyranny of the minority.
4. Policy decisions in consociational democracies need considerable time due to processes of inclusion and negotiation. Hence, these forms of government may appear particularly disadvantageous if the political system has to respond quickly to threats, such as an imminent economic crisis.
5. Negotiations and compromise imply that policy decisions have a considerable likelihood of being incoherent and hence less efficient than policy packages that are shaped by a single political goal or ideology.
6. This system of compromise is vulnerable to populist movements on the right and the left, and it may frustrate citizens in their political activity.

These disadvantages used to be considered as the price for a stable democracy in deeply divided societies. However, recent empirical research found other benefits and evidence that the disadvantages of consociatonal systems are much lower than suggested by the critics. In addition to political stability, consociational democracy is inclusive, giving smaller groups (or their representatives) a much better say in the political process. This may outweigh some of the democratic deficits of consociationalism. In addition, since the policy packages are based on compromise, implementation of policies may be quicker and more efficient than in competitive democracies. Hence, consensus systems may make up some of the loss in time and efficiency during the phase of policy formation. Finally, consensus democracies tend to be less violent, economically superior, and "kinder and gentler" compared with majoritarian democracies. In consensus systems, citizens are politically more satisfied; the government is more redistributive; strike activity, unemployment, inflation, and inequality are less; political activity and participation are higher; and policies are more "woman friendly." Some of these findings are controversial; however, there is strong evidence that the political and economic achievements of consociational democracies are not inferior to those of majoritarian democracies.

Criticism

Criticism of consociational democracy pointed to conceptual and empirical deficits. While the literature on consociationalism argued that this decision mode allows for a stable political order in deeply divided societies, Ian Lustick (1979) developed a control approach that focuses "on the emergence and maintenance of a relationship in which the superior power of one segment is mobilized to enforce stability by constraining the political actions and opportunities of another segment" (p. 328). Brian Barry (1975) detected a tautological argument in Lijphart's work in that consociationalism was described as a deliberate choice of elites, since for this choice elites need to compromise from the very beginning. Other authors emphasized that separated societal segments as the sociocultural precondition of consociational democracies started to decline exactly when consociational techniques were strongly used in the 1950s and 1960s. Finally, critics identified problems in mixing normative and empirical typologies, noting that Lijphart's list of favorable conditions for consociationalism changed since the first publications about consociationalism.

Rein Taagepera (2003) argues that Lijphart's indicators of the first dimension of consensus democracy are mainly output variables, which are not directly amenable to institutional design, while the indicators of the second dimension are based on expert judgments and measured mostly on a nominal level. Likewise, Steffen Ganghof (2005) states that Lijphart's concepts refer to institutions, while many of his indices relate to observed behavior such as coalition building or social partnership.

Two criticisms have strong implications for the political importance of consociational democracy. Lijphart suggests that elites have an option for consociational democracy and it is up to them to decide in favor or against. Based on this assumption, he recommends to politicians in deeply divided societies to adopt the consensus or, even better, the consociational system. Already early in the debate, this voluntaristic approach was questioned by Lehmbruch. If consociationalism depends on learning processes, long-term elite socialization, and supporting institutions, it cannot be inserted successfully from outside into a political system that lacks these enabling conditions.

The second problem may be the decline of the prerequisite of any system of compromise in segmented societies: Individuals are integrated into sociocultural segments, and these segments have legitimate political representatives. The processes of individualization and modernization tend to undermine this precondition. Increasingly—even in formerly strongly segmented societies—individuals do not build stable links to groups and political organizations. Hence, the sociocultural foundations of consociationalism vanish, endangering the long-term stability of consensus institutions. A limited decline of consociational democracy has been observed in many countries (see Adrian Vatter, 2008). On the other hand, one of the preconditions of an ideal type of majoritarian democracy suffers from the same modernization processes. Modern societies do not consist any longer of two groups of similar size whose political representatives alternate in government. Rather, political integration in large groups declines, and democracies may benefit from institutions of negotiation that are able to include these numerous small groups of modern society, none of which is able to produce a political majority. The obvious problem seems to be that these societies may be able to solve their problems with consociational techniques; however, they lack the sociocultural preconditions to use these techniques.

Klaus Armingeon
University of Bern
Bern, Switzerland

See also Democracy, Quality; Democracy, Theories of; Democracy, Types of

Further Readings

Armingeon, K. (2002). The effects of negotiation democracy: A comparative analysis. *European Journal of Political Research, 41,* 81–105.

Barry, B. (1975). Political accommodation and consociational democracy. *British Journal of Political Science, 5,* 477–505.

Birchfeld, V., & Crepaz, M. M. L. (1998). The impact of constitutional structures and collective and competitive veto points on income inequality in industrialized democracies. *European Journal of Political Research, 34,* 175–200.

Bogaards, M. (2000). The uneasy relationship between empirical and normative types in consociational theory. *Journal of Theoretical Politics, 12*(4), 395–423.

Ganghof, S. (2005). Normative modell, institutionelle typen und beobachtbare verhaltensmuster: Ein vorschlag zum vergleich parlamentarischer demokratien [Normative models, institutional types, and observable pattern of behavior: A proposal for the comparison of parliamentary democracies]. *Politische Vierteljahresschrift, 46,* 406–431.

Kaiser, A. (1997). Types of democracy: From classical to new institutionalism. *Journal of Theoretical Politics, 9,* 419–444.

Lehmbruch, G. (1967). *Proporzdemokratie: Politisches system und politische kultur in der Schweiz und Österreich* [Proportional democracy: Political system and political culture in Switzerland and Austria]. Tübingen, Germany: Mohr (Siebeck).

Lehmbruch, G. (1993). Consociational democracy and corporatism in Switzerland. *Publius: The Journal of Federalism, 23,* 43–60.

Lijphart, A. (1968). *The politics of accommodation: Pluralism and democracy in the Netherlands*. Berkeley: University of California Press.

Lijphart, A. (1999). *Patterns of democracy: Government form and performance in thirty-six countries*. New Haven, CT: Yale University Press.

Lijphart, A. (2008). *Thinking about democracy. Power sharing and majority rule in theory and practice*. New York: Routledge.

Lustick, I. (1979). Stability in deeply divided societies: Consociationalism versus control. *World Politics, 31,* 325–344.

Schmidt, M. G. (2008). *Demokratietheorien* [Democratic theory] (4th ed.). Wiesbaden, Germany: Verlag für Sozialwissenschaften.

Taagepera, R. (2003). Arend Lijphart's dimensions of democracy: Logical connections and institutional design. *Political Studies, 51,* 1–19.

Vatter, A. (2008). Swiss consensus democracy in transition: A re-analysis of Lijphart's concept of democracy for Switzerland from 1997 to 2007. *World Political Science Review, 4*(2). Retrieved October 12, 2010, from http://www.bepress.com/wpsr/vol4/iss2/art1

Democracy, Direct

Direct democracy means forms of direct participation of citizens in democratic decision making in contrast to indirect or representative democracy, based on the sovereignty of the people. This can happen in the form of an assembly democracy or by initiative and referendum with ballot voting, with direct voting on issues instead of for candidates or parties. Sometimes the term is also used for electing representatives in a direct vote as opposed to indirect elections (by voting for an electing body, electoral college, etc.), as well as for recalling elected officeholders. Direct democracy may be understood as a full-scale system of political institutions, but in modern times, it means most often specific decision-making institutions in the broader system environment of representative democracy. The following sections (a) introduce some historical background and theoretical ideas, (b) elaborate on various forms of direct democracy, (c) describe regulations and use of direct democracy in some important countries, and (d) present main issues and controversies.

History and Theoretical Context

The most important historical reference of direct democracy is to assembly democracy in ancient Greek city-states, particularly Athens, where decisions were taken by people's assemblies of some 1,000 male citizens. Later, people's assemblies were used in many Swiss cantons and towns, as well as in town meetings in some American colonies and states. Early U.S. states also started using procedures in which constitutions or constitutional amendments were ratified by referendums, which later became common in the United States. Popular sovereignty, proclaimed in the French Revolution, had rather been distorted, however, in Napoleon's autocratic plebiscites. Switzerland and many U.S. states incorporated direct democracy in their constitutions during the 19th century, while Germany and few other countries adopted some elements after World War I. In a more general perspective, the ensuing introduction or practical use of direct-democratic institutions originated from three major types of developments:

1. social class conflict to curb the political power of a dominating oligarchy (e.g., Switzerland, U.S. states);
2. processes toward political/territorial autonomy or independence for legitimizing and integrating the new state unit (beginning after World War I); and
3. processes of democratic transformation from authoritarian rule (e.g., Germany's regional states after 1945, some Latin American countries).

Some countries show gradual reform developments (e.g., Uruguay).

Modern democracy most often developed not from the starting point of assembly democracy but, under absolutist or feudal conditions, from people gradually claiming a larger share of political representation and extension of representative voting rights. Constitutions, civil rights, and universal suffrage, which had been achieved in European and many other countries (generally by the end of World War I), were usually identified with "democracy" on the normative basis of the principles of popular sovereignty, freedom, and political equality. Thus, in many countries and theories, these principles have been tied to and absorbed by a narrow notion of representative democracy rather than being used to support a more comprehensive concept of democracy.

Normative theory of direct democracy still rests basically on popular sovereignty, freedom, and political equality, with Jean-Jacques Rousseau as the outstanding theorist of unanimous consent of the people for a free republican constitution and subsequent forms of participation. During the 19th century, these principles were increasingly challenged, or they were deprived of their substance beyond representative institutions. So, in many countries, direct-democracy institutions have not been established or implemented since representative elites developed a strong interest in monopolizing power. In addition, pragmatic theories contended that direct democracy could not work under space and time conditions of large modern states.

With this background of historical and theoretical restrictions, the normative theory of direct democracy cannot exclusively rest on popular sovereignty, which is also claimed by representative democracy. More specific arguments originate from the participatory theory of democracy and the critique of a lack of responsiveness and legitimacy of representative (party) democracy. The two sets of democratic institutions are distinguished by basic features of direct participation: (1) direct democracy focuses on specific *issues*, in contrast to voting on candidates and general programs for long terms of office, and (2) citizens themselves act as *decision makers* rather than delegating these powers. Like electoral systems, a variety of procedural forms, designs, and regulations are likely to influence processes and outcome. One must also keep in mind that direct-democratic processes cannot operate in isolation but are always linked to the structures of an overall political system that includes major representative institutions. Thus, interactions between the two types of institutions will be an important challenge for analysis. For instance, as George Tsebelis notes, referendum voters can be seen as an additional veto player. Some authors contend that direct democracy may undermine representative democracy, while others focus on the deliberative functions for a democratic public sphere and the capacity for integrating citizens in the democratic process. One can also assume that basic types or forms of direct-democracy procedures may result in different consequences.

Variety of Forms and System Environment

Direct democracy comes in a variety of institutional forms, with the common feature of procedures focusing on popular votes on political issues. Their main forms can be distinguished by the actors who start the procedure: *Mandatory referendums* have to be held when a referendum vote is required by law (e.g., a constitution) for deciding a specific subject. *Referendums of governmental authorities* take place when a president, cabinet, or legislature decides, under preregulated conditions or ad hoc, to call a popular vote on a particular issue. Sometimes, a minority of a legislature also is entitled to demand such a vote. *Citizens' initiatives* that are supported by a required number of signatures allow the electorate to vote on political measures proposed by a group, on legislative acts by a parliament not yet in force, or on existing laws (citizen-demanded referendums). A popular vote may be binding according to the simple or specific majority or turnout requirements for a valid vote or may be defined as only consultative or advisory.

Some jurisdictions provide an agenda initiative that allows citizens with the support of a minimum number of signatures to place a particular issue on the agenda of a government or legislative authority. Such proposals have to be considered by the authority addressed, but they do not lead to a referendum vote.

There is some ambiguity and controversy as to whether procedures with a focus on directly electing or recalling holders of public office (executive positions, legislators) may be meaningfully included in the concept of direct democracy. These procedures refer, in fact, to the institutional system of representative democracy and its typical processes and, therefore, are not at the core of debates on direct democracy. However, there may be some differences in the degree to which voters have a direct influence on the final outcome of an electoral procedure (e.g., fixed or flexible list of candidates, direct vote, or vote for members of an intermediate body). In recall procedures, interrupting routine patterns of fixed office terms may stress the aspect of citizens reclaiming control of office functions. In practice, recall options of executive office holders are much more common than of members of legislative bodies or of complete legislatures.

Procedural types of direct democracy should be distinguished according to the main initiating actor of a procedure because they typically show different features regarding the agenda setter, the contents and wording of the proposal, the function of the ballot vote in terms of legitimation, innovation, and so on.

Governmental authorities initiating a referendum vote generally seek legitimation for policies on the government agenda, will regularly advocate an affirmative vote, and will have many ways of influencing process and outcome, including official communication resources. Therefore, the term *plebiscite* is often used, even more so when they are employed by autocratic or dictatorial regimes that cannot be called democratic at all.

Mandatory referendums also very often originate from governmental authorities entitled to bring forward proposals for which ratification by a referendum vote is required, particularly in the case of constitutional amendments or matters of state sovereignty, territory, or identity. Thus, whereas a popular vote on such specific subjects is required by law, the agenda and the substance of the referendum proposal are most often determined by governmental authorities. In some jurisdictions, however, specific issues, again like constitutional amendments, may also be proposed by citizens' initiative and lead to a mandatory ballot vote (Switzerland, the United States, or German states).

In citizens' initiative procedures, the agenda for issues and the proposals generally originate "bottom up" from some opposition or civil society groups that demand new political measures or legislation (law-promoting initiative) or object to a particular government project or legislative act (law-controlling initiative). In such a setting, the political initiative comes from social or minority forces, whereas governmental authorities are likely to be in a defensive position and want to defeat the proposals in a referendum vote.

Except for ad hoc referendum calls by governmental authorities, procedures of direct democracy, particularly citizen-initiated procedures, are regulated in various aspects. The area of admissible subject matters may be very restrictive, the number of signatures required for qualifying an initiative for a ballot vote may range from about 1% to one third of eligible voters, the time allowed for collecting signatures may be very short. Requirements for the validity of a popular vote may also vary from a majority of voters to qualified or double majorities or to specific turnout quorums. Usage will clearly be restrained by high initiating or validity requirements and initiating actors with strong resources will be privileged. Yet a higher level of approval may support the legitimacy of a vote.

Countries and Developmental Background

Provisions for direct-democratic instruments as well as their usage are distributed rather unevenly across continents, countries, and different levels of states.

National Level

On the national level, procedures and usage are most frequent in Europe and Latin America, whereas in Africa, Asia, or North America their number is small. Switzerland traditionally has the most elaborated system of direct democracy at the national, cantonal, and municipal levels. On the national level, mandatory referendums on constitutional amendments were introduced in 1848, citizen-demanded rejective referendums on new legislation of parliament in 1874, citizens' initiatives on constitutional amendments in 1891, and mandatory referendums on major international treaties in 1921. Up to 2008, 222 mandatory referendums on constitutional and treaty issues, 162 rejective ("facultative") referendums, and 165 citizen-initiated referendums on constitutional amendments have been held. On the lower levels, even more instruments are often available, such as the mandatory financial referendum and legislative citizens' initiative. More than one of the instruments of direct democracy are also provided for and practiced on the national level in Uruguay and, more recently, in some of the Eastern European countries such as Latvia, Lithuania, or Hungary.

On the national level in most other countries, basically one specific instrument stands out. In Italy, for instance, most of the popular votes originated from citizens' initiatives for an "abrogative referendum" to repeal an existing law or parts thereof (some 60 cases since 1970). In Australia and Ireland, only mandatory referendums on constitutional amendments are possible. In Austria, the national parliament can call a vote (used only once). In the Fifth Republic of France, the president has the right to call a referendum on important matters of sovereignty and state structure since 2008; one fifth of the National Assembly in combination with 10% of the electorate can also do so. In many countries, governmental authorities call referenda not on a preregulated but on an ad hoc basis. The subjects of these referenda have included accession to the European Economic Community or to the European Union (EU), independence or national unity in Canada, and ratification of a new constitution or of an agreement for conflict or constitutional settlement in Bolivia, Chile, Kenya, Russia, Spain, and South Africa.

Regional Level

The level of regional states within federations also offers important examples of direct democracy. In particular, about half of the states in the United States provide citizens' initiative rights and some other procedures that quite often allow referendum votes on public finance. In many of these states, including Oregon, California, Colorado, and North Dakota, citizens' initiatives (introduced before 1914) are frequently used. In Germany, all regional states (*Länder*) have introduced citizens' initiatives, which, on the average, are used less often due to more restrictive requirements. In the United States and in Germany, initiatives and referendums are available and frequently used on the municipal level as well. In some countries, including the Czech Republic, Japan, Norway, and Poland, initiatives and referendums are only possible at the municipal level.

Paths to Direct Democracy

In systems of representative democracy, the introduction of instruments of direct democracy is not very likely since they are regarded as undermining established power structures. The historical origins of direct democracy institutions can be distinguished in typical paths:

1. antioligarchic conflict, where initiative and referendum are directed against economic-political domination, like in Switzerland and the progressive movement in the Western states of the United States before 1914;
2. system transformation toward democracy, like in post-Fascist Italy (implementation delayed until 1970), in German regional states after dictatorships in 1945 and 1990, or in Latin America in the 1980s and 1990s;
3. state independence, like in Ireland (1922/1937) or in Latvia, Lithuania, and Slovenia after Soviet or Yugoslav domination.

From the first two paths, initiative instruments with lower requirements quite often emerged, whereas national independence often coincided with requirements stressing majoritarian unity. The origins of referendums called by governmental authorities are much more dispersed since they are mainly regarded as under government control and, thus, less dangerous to the governmental system. Issues of state independence, but also accession of states to supranational organizations such as the EU, very often are dealt with in government-initiated or in mandatory referendums. The first instrument of direct democracy with a transnational character is the European Citizens' Initiative of the Lisbon Treaty, entered into force on December 1, 2009 (Article 11.4), an agenda initiative that allows 1 million European citizens to propose legislation to the European Commission (without a referendum).

Issues and Controversies

Discussions on direct democracy institutions deal with several issues. The strongest normative grounds for direct democracy are the democratic principles of popular sovereignty, political equality, and all the arguments for participative democracy that support the idea that all citizens should have the right not only to elect representatives but also to vote on policy issues in referendums. Since assembly democracy cannot be an option in modern societies, direct democracy institutions are not regarded as a full-scale system alternative to representative democracy but as a supplement to or counterweight within democratic systems with major representative features. Nevertheless, the institutional difference and competition between representative and direct democratic processes lie at the core of the controversy whether direct democracy contributes to undermining representative democracy or can offer enrichments of democracy.

In general, representative democracy is often seen as superior because general elections give citizens an encompassing choice between alternative governments and complex and coherent programs, because governments and parliaments have greater capacity for informed decisions including expert judgment, and because representatives can be held accountable for their decisions. Arguments in support of direct democratic instruments refer to various aspects:

1. Direct democratic issue voting can, during terms of office, deal with issues that have not been discussed at general elections. Citizens' initiatives in particular can enrich the political agenda and, thus, contribute to the function of political articulation and innovation. The range of political actors tends to be broader than already present in the party system.

2. Direct democracy also offers citizens additional and more specific instruments of political control during terms of office, particularly initiative proposals and citizen-demanded referendums to reject new legislation or delete existing laws.

3. One major area of controversy deals with information, competence, and the quality of decision making. While representative institutions may indeed hold intense deliberations on many subjects, direct-democratic decision-making processes can also provide for specific issues the opportunity of intense and widespread public debates, during which citizens can become informed about controversial value and factual considerations. Yet as voters are often described as badly informed and incompetent, the danger of manipulation by resourceful actors (parties, strong interest organizations, corporations, media) is a major issue. Design and regulations, however, can make a difference, for example, comparing Switzerland and U.S. states such as California, since Switzerland does not allow TV advertising in referendum campaigns. Hanspeter Kriesi (2005) and Daniel Smith and Caroline Tolbert (2004) especially found that, as a general trend, referendum debates and campaigns provide a major potential for dissemination of information and for political education. Important factors are a broader field of political actors in such campaigns and more intense communication of arguments in the media.

4. How voters select their choice in referendum voting attracted two rather opposite assumptions. According to one, party orientations might be simply duplicated in issue voting behavior; the other one contends that interest groups, media, and even "demagogues" can influence voters strongly. The trend in the literature seems to be that the less informed parts of the electorate look for party orientation, while voters who are better informed and educated may take a more independent choice by using more arguments for forming opinions. Thus, as Ian Budge (1996) notes, direct democracy must not be unmediated, since governments and parties can also play an important mediating role. This may be particularly true when government authorities initiate a referendum vote in optional or mandatory referendums.

5. One promise of direct democracy is that more political participation can be realized. This is surely the case since more opportunities and occasions to debate policy issues and to vote in referendums are offered. Nevertheless, some criticism remains that the participation goal is not realized, particularly for social groups that also participate little in electoral politics. It is argued that in referendums turnout is often lower than in general elections and that referendums lead to a lower turnout in elections also. Yet this cannot be generalized since turnout varies significantly dependent on issues, for example, in Switzerland from 30% to around 80% of registered voters. In addition, when ballot votes regularly do not take place in conjunction with general elections (like in Switzerland), turnout variation will be stronger than when referendum votes are mostly held on election days (like in the United States). Other shortcomings of extended participation of direct democracy are seen in a "social bias," where lower social strata with deficits in status, income, and education tend to be less motivated or competent to participate in discussions or in voting. Again, at least to some degree, this seems to be balanced by policy issues specifically relevant for these groups in which they participate to a larger degree.

Finally, if a larger share of citizens do not participate, but abstain from voting, legitimation problems in referendum votes may arise. In some jurisdictions, regulations respond by requiring a qualified majority for a valid vote in the form of turnout or approval quorums. The disadvantage of turnout quorums is that, in turn, it invites even more abstentions and campaigns to abstain even from voting "no," while approval quorums at least devaluate majority votes.

6. Direct democracy institutions have also been reflected in their relation to majorities and minorities. Citizen-initiated procedures are supposed to serve as potential instruments of minorities since they can present new proposals or demand a referendum on new legislation. This is likely to be true for placing issues on the agenda. In the referendum vote, however, the majority principle applies, which means that minority rights or interests can be endangered. Specific concerns relate to basic rights of minorities, which, however, can best be protected against offensive majority rule by constitutional guarantees, courts, and prereferendum constitutionality checks. Sometimes, the validity of referendum votes is regulated by qualified or double majorities to protect minorities. More generally, referendums can also support developments toward autonomy or even independence of regional "minority" populations.

7. Policy impacts of direct democratic decision making also received attention. Research on economic and financial effects at the regional and local levels of Swiss and U.S. direct democratic institutions found beneficial consequences in macroeconomic and fiscal performance. According to one thesis, strong interest groups will gain more from initiatives and referendums, whereas empirical economic studies, such as the one by John Matsusaka (2004), tend to find advantages rather with the broader population. In other policy issues such as the environment or moral topics, tentative and controversial evidence prevails.

8. Other consequences attributed to direct democracy refer to structural or system impacts on representative democracy or the overall system of democracy. In the case of Switzerland, particularly, it has been argued that direct democracy had a long-term effect toward a system of consensus democracy as opposed to majoritarian democracy. Mechanisms of consensus governments may indeed have developed to anticipate and integrate as many interests as possible, which otherwise might be able to initiate referendums against new legislation (instruments such as citizens' initiatives and mandatory referendums would be less relevant here). In other jurisdictions, however (e.g., in Italy or German or U.S. states), similar effects away from majoritarian party competition and toward consensus democracy could not be observed; in presidential systems such as in the U.S. states legislative majorities and the executive are disconnected anyway. Thus, generalizations from the Swiss example on a developmental logic toward consensus government should not be easily drawn.

One should keep in mind that government-initiated and government-controlled referendums may in many respects show distinct features from citizen-initiated procedures. Government-driven instruments tend to be more influenced by policy projects and campaign capacities of central political authorities. Citizen-initiated procedures are more open for minorities, participation, innovation, and government control, yet they are less likely to succeed in the ballot vote. Nevertheless, as a process, they tend to offer a greater potential for supplementing and balancing the institutional shortcomings and power structures of representative democracy. Particularly in times of political crises, direct democracy can provide an important function in offering channels for reactivating popular sovereignty as the fundamental value and force of democracy. This power of preserving the sources of popular sovereignty alone makes it worth while to keep direct democracy going under routine conditions of democracy.

Theo Schiller
Philipps-Universität Marburg
Marburg, Germany

See also Democracy, Quality; Democracy, Theories of; Referenda; Representation

Further Readings

Altman, D. (in press). *Power to the people? A global view on direct democracy.* Santiago, Chile: Instituto de Ciencia Política–Pontificia Universidad Catolica de Chile.

Budge, I. (1996). *The new challenge of direct democracy.* Cambridge, UK: Polity Press.

Butler, D., & Ranney, A. (Eds.). (1994). *Referendums around the world: The growing use of direct democracy.* Basingstoke, UK: Macmillan.

Initiative and Referendum Institute Europe. (2008). *Guidebook to direct democracy in Switzerland and beyond*. Marburg, Germany: Author.
Institute for Democracy and Electoral Assistance. (2008). *Direct democracy: The international IDEA handbook, 2008*. Stockholm: Author.
Kriesi, H. (2005). *Direct democratic choice: The Swiss experience*. Lanham, MD: Lexington Books.
Leduc, L. (2003). *The politics of direct democracy: Referendums in global perspective*. Peterborough, ON, Canada: Broadview Press.
Matsusaka, J. G. (2004). *For the many or the few: How the initiative process changes American government*. Chicago: University of Chicago Press.
Setälä, M., & Schiller, T. (Eds.). (2009). *Referendums and representative democracy: Responsiveness, accountability and deliberation*. London: Routledge.
Smith, D. A., & Tolbert, C. J. (Eds.). (2004). *Educated by initiative: The effects of direct democracy on citizens and political organizations in the American states*. Ann Arbor: University of Michigan Press.

Democracy, Quality

In recent decades, discontent, dissatisfaction, and alienation have been growing political phenomena in old, established democracies. At the same time, more and more countries have been defined as democratic. Democratization has also brought about a large number of so-called hybrid regimes. From these three different perspectives, the same question emerges: What is the quality of democracy in any specific country, and what does this tell us in terms of (a) how actual and deep the satisfaction or dissatisfaction is, (b) what is behind democratic appearances, and (c) how ambiguous cases can be turned into democracies by improving their quality? This entry focuses on the notion of quality, its key dimensions or qualities, and the mechanism of quality subversion as a key aspect to explore when dealing with this topic.

What Is Quality?

There are at least three main streams of literature dealing with this topic. A first group of scholars have been working on democratizations, consolidation, and crisis and in their analyses of democracy have also pointed out aspects related, directly or indirectly, to the topic. Among them are Guillermo O'Donnell with his notion of "delegative democracy," Arend Lijphart with his thesis on the superiority of consensual democracy vis-à-vis majoritarian democracies in terms of implementing democratic quality, and David Altman and Aníbal Perez-Liñan who refer to three aspects that draw on Robert Dahl's concept of polyarchy (civil rights, participation, and competition). The second group of scholars focused on established democracies, especially those belonging to the Anglo-Saxon tradition—that is, the United Kingdom, Canada, and Australia. Among them, David Beetham (see especially David Beetham, Sarah Bracking, Iain Kearton, & Stuart Weir, 2002) developed an analysis in terms of audit—that is, a systematic assessment of institutional performance against agreed criteria and standards. The auditing procedure should follow four steps:

1. Identify appropriate criteria for assessment.
2. Determine standards of good or best practice that provide a benchmark for the assessment.
3. Assemble the relevant evidence from both formal rules and informal practices.
4. Review the evidence in the light of the audit criteria and defined standards to reach a systematic assessment.

A number of authors followed Beetham by implementing his framework in other countries, such as Canada and Australia (see, e.g., Marian Sawer, Norman Abjorensen, & Phil Larkin, 2009; see also Todd Landman, 2006).

The third stream is formed by a number of data banks, such as those of Polity IV, Freedom House, the Economist Intelligence Unit, World Bank, and also the Bertelsmann Index, which provide measures of aspects related to democratic performance and, more generally, to the quality of democracy. The different data banks are massive efforts to provide scores and rank orders of a large number of countries or, in some cases (first and foremost Freedom House), of all existing independent countries on key aspects such as rule of law and freedom.

Thus, phenomena such as democratization processes and political science research paved the way for the development of a more systematic theoretical analysis of the quality of democracy, where a key point is spelling out a clear notion of quality. A survey of the use of the term in the industrial and marketing sectors suggests three different meanings of *quality*:

1. Quality is defined by the established procedural aspects associated with each product; a "quality" product is the result of an exact, controlled process carried out according to precise, recurring methods and timing; here the emphasis is on the *procedure*.
2. Quality is defined by the structural characteristics of a product, be it the design, materials, or functioning of the good or other details that it features; here, the emphasis is on the *content*.
3. The quality of a product or service is indirectly derived from the satisfaction expressed by the customers, by their repeated requests for the same product or service, regardless of how it is produced or what the actual contents are or how the consumers go about acquiring the product or service; according to such a meaning, the quality is simply based on *result*.

The three different notions of quality are grounded in procedures, contents, or results. Each has different implications for empirical research. Importantly, even with all the adjustments demanded by the complexity of the object under examination—democracy—it is still necessary to keep these conceptualizations of quality in mind as definitions and models of democratic quality/ies are elaborated. The next questions, then, are "What is a quality democracy?" and, more precisely, "What are the procedural, content, and result qualities of a democracy?"

"Good" Democracy

A quality democracy or a "good" democracy presents a stable institutional structure that realizes the liberty and equality of citizens through the legitimate and correct functioning of its institutions and mechanisms—that is, a good democracy is first and foremost a broadly legitimated regime that completely satisfies citizens (*quality in terms of result*); one in which the citizens, associations, and communities of which it is composed enjoy liberty and equality, even in different forms and degrees (*quality in terms of content*); and one in which the citizens themselves have the power to check and evaluate whether the government pursues the objectives of liberty and equality according to the rule of law (*quality in terms of procedure*).

Alternative normative definitions of democracy could be recalled, such as liberal democracy, responsive democracy, participatory democracy, deliberative democracy, associative democracy, egalitarian or social democracy, and good governance. However, regardless of what normative definition of democracy is accepted, empirical analysis tends to reveal the same specific features.

What Qualities?

Eight possible dimensions or *qualities* on which a "good" democracy might vary are at the core of the empirical analysis to cover the normative notions of democracy mentioned above. The first five are procedural dimensions. Though also relevant to the contents, these dimensions mainly concern the rules. The first procedural quality is the *rule of law*. The second and third procedural qualities regard the two forms of *accountability* (*electoral* and *interinstitutional*). The fourth and fifth are the classic *participation* and *competition*, which, however, have a special theoretical status (see below). The sixth and seventh dimensions are substantive in nature and deal with quality as defined in terms of content: The sixth is full respect for rights that are expanded through the achievement of a range of *freedoms*. The seventh is the progressive implementation of greater political, social, and economic *equality*. The final, eighth, dimension concerns the *responsiveness* or correspondence of the system to the desires of the citizens and civil society in general, and this is quality defined in terms of results. Some essential considerations about these dimensions have still to be added.

The institutions and mechanisms of *representative* democracies are actually the main objects of the empirical analysis of democratic qualities. This is not to ignore the direct democracy as the highest expression of democratic quality but to acknowledge the centuries-long empirical experience of representative democracies. If the analysis is to focus on representative democracies, then accountability—a core feature in the representative democracy—becomes a truly central dimension insomuch as it grants citizens and civil society in general an effective means of control over political institutions. This feature attenuates the difficulties that exist objectively when there is a shift from direct to representative democracy.

The principal actors of such a democracy are the citizen-individuals, the territorial communities, and the various forms of associations with common values, traditions, or aims. In this sense, the possibility for good democracy exists not only in the case of a defined territory with a specific population controlled by state institutions under a democratic government but also for wider ranging entities. The main point is that the above-named subjects are at the heart of a democracy in which the most important processes are those that work from the bottom up, and not vice versa. In this way, the transfer of the analytical dimensions from the national level to the supranational level—though not uncomplicated and without difficulty—is possible.

It is particularly important to point out the specific empirical aspects to explore for each of the eight *qualities*. They are briefly summed up here.

Quality in Terms of Procedure

Rule of Law

If the rule of law is at stake, the relevant elements to examine are

1. individual security and civil order, with a focus on right to life, freedom from fear and torture, personal security, and right to own property guaranteed and protected through the country;
2. an independent judiciary and a modern justice system, focused on mechanisms establishing an independent, professional, and efficient judiciary system that allows equal access to justice, free of the undue pressures and enforcement of decisions;
3. institutional and administrative capacity to formulate, implement, and enforce the law, where the focus is on the governance system (parliament and government), the capability to ensure the production of quality legislation and the implementation through the country of a transparent policy-making process allowing for the participation of civil society, and the presence of a professional, neutral, accountable, and efficient state bureaucracy;
4. effective fight against corruption, illegality, and abuse of power by state agencies, where the focus is on the existence and implementation of the comprehensive legislative framework to prevent and fight the corruption; and
5. security forces that are respectful of citizen rights and are under civilian control; the focus is on the mechanisms of civilian control over security forces as well as on efficient, noncorrupted, disciplined police forces respectful of human and political rights.

Electoral Accountability

If the quality to analyze is electoral accountability, the best strategy for detecting it empirically is to refer to its most immediate conditions:

1. free, fair, and recurrent elections, with their specific procedural aspects;
2. plural and independent information;
3. freedom of the party organization and related aspects; and
4. the presence and stability of alternatives.

Interinstitutional Accountability

If the quality to analyze is interinstitutional accountability, the main subdimensions and indicators to explore refer to

1. legislative–executive relations, with special focus on the parliamentary opposition or the role of the legislative body;
2. constitutional courts;
3. ombudsmen;
4. audit courts; and
5. modes and extent of decentralization.

Political Participation

If political participation has to be considered, one should look at

1. identifying participation, to become a "part," to revive, or to restate the belonging or identification with a group of a different sort and
2. instrumental participation, to try to achieve some goal.

So the basic forms in the case of conventional participation are referendum, electoral participation at the country and local levels, and membership and other forms of affiliation in political organizations and interest associations. In the case of nonconventional participation, the basic forms are strikes, demonstrations, and riots; there are other forms involving participation, in addition to forms of participation with regard to specific policies.

Political Competition

The basic salient subdimensions of political competition are the following:

1. competition among political and societal actors, characterized by freedom for all political parties to compete with each other and by fairness of political competition;
2. competition within political and societal actors; and
3. the output side of political competition.

Quality in Terms of Content

Freedom

The three main subdimensions of freedom are the following:

1. personal dignity,
2. civil rights, and
3. political rights.

For all of them, there is the existence of opportunity in the legal system of the country and the actual guarantee of each.

Equality

The guarantee and effective implementation of social, economic, and cultural rights helps reduce differences in those domains. In addition, when considering equality, other basic aspects to analyze include the following:

1. the allocation of economic resources within the population,
2. the extent of poverty,
3. the diffusion of education,
4. the existence of gender discrimination, and
5. the existence of ethnic discrimination.

Quality in Terms of Result

Responsiveness

Responsiveness refers to the legitimacy of government—that is, the citizens' perception of responsiveness. In other words, the empirical aspects to consider should be the diffusion of attitudes favorable to the existing democratic institutions and the approval of their activities because of the assumption that attitudes of satisfaction show the effective perception by civil society of existing responsiveness. On this quality, limited resources and economic constraints on public spending affect the responsiveness of even the wealthiest countries. Likewise, the persistent problems posed by unemployment and immigration are also illustrative of the near impossibility of finding generally satisfactory, legitimate, and responsive solutions in contemporary democracies. Indeed, the situation is more and more characterized by discontent, dissatisfaction, fear of poverty, and general democratic malaise.

Recurrent Patterns of Subversion

A different, meaningful perspective on the analysis of qualities, above all a more realistic one, is to look at all the recurrent ways in which elites and citizens consciously or otherwise try to subvert those qualities for their political or private purposes. Here the eight qualities with their possible and often probable subversions are reviewed.

Rule of Law

Starting with the rule of law, first and foremost, a rigorous application of laws or, in certain cases, the relationships with an only apparently efficient bureaucracy can have particularly negative consequences for the most socially weak and vulnerable members of society. Then, there is the possible use of the law as a genuine political "weapon." Here, one can see a persistent and diffuse temptation for politicians to use the law against their adversaries if, for example, the opposition is condemned to remain as such for a long time and has no chance of electoral victory in the near future. Politicians are also tempted to use judicial acts to reinforce their own position against the opposition. In other cases, when there is collusion among politicians, the judges themselves, with the support of the media, are tempted to turn to the judiciary in retaliation for certain political decisions that they consider unacceptable. On a different level, there is also a growing tendency among individual citizens or economic groups to resort to the law to assert their own interests. Some scholars note this phenomenon as a "juridification" of contemporary democracy. Finally, and not altogether different, there is the popular and diffuse cultural attitude that interprets the law as a severe impediment to realizing one's own interests that should be circumvented in any way possible. This attitude, which is common in various countries throughout the world from Southern Europe to Latin America, Eastern Europe, and also Asian democratic countries, extends from the popular to the entrepreneurial classes.

Accountability

With respect to electoral accountability, given the well-known opacity of political processes and the complexity of reality, politicians have ample opportunity to manipulate their contexts in such a way as to absolve themselves of any concrete responsibility. Accountability frequently becomes a catchphrase more connected to the image of a politician than to any decisions he or she may have taken or results he or she might have produced. Negative outcomes are easily justified by making reference to unforeseen events or by taking advantage of a favorable press to influence public opinion. At the same time, good results, obtained sometimes at the cost of sacrifices by the governed, might result in negative or punitive judgments for the governor at the time of the next elections. Thus, when a politician supports something that is unpopular but necessary for the welfare of the state, he may lose the support of the electorate. The very action, often ideological and instrumental, of parties or other components of the political opposition or even of media actors in a position to conduct public processes can, sometimes on inconsistent grounds, make the effective implementation of electoral accountability more difficult. The lack of clear distinctions between incumbent leaders and party leaders—the head of government often also controls the parties—means that parties, be they of the opposition or of the majority, are hindered in carrying out their role as watchdogs for their constituents. At the parliamentary level, party discipline is considered more important than accountability toward the electors and, in practice, the parliamentary majority supports the government without controlling it. Furthermore, there should also be a clear distinction between the responsible leader, either of the government or of the opposition, and the intermediate layers of party actors that range from militants to sympathizers. These latter should trigger a bottom-up process that gives direction to how parties should control the government or organize their opposition. However, recent empirical studies on party organization in a number of advanced democracies indicate an opposite trend, which is characterized by strong party leaders who act in collusion (instead of in competition) with other parties or party leaders. The most extreme scenario relating to this phenomenon is that parties, supported by public financing, effectively form "cartels" where the political opposition is actually disappeared.

Citizens in European countries encounter further difficulties in ensuring electoral accountability because of the existence of the supranational dimension created by the European Union (EU). The most fitting example of how governments in these countries avoid accountability is the well-known tactic of blame shifting. Here, the political responsibility for every unpopular decision taken by the government is shifted from the national to the European level, even if they concern clear-cut issues such as streamlining national administrations or reorganizing state finances to meet large national deficits. Governments or national politicians justify actions resulting in widespread public opposition by claiming that their hands were forced by opposing coalitions in the Council of Ministers of the EU or in the European Council of prime ministers and chiefs of state or by votes in the European Parliament.

As José Maravall (1997) has discussed, there are many ways by which government leaders can avoid electoral accountability. At the same time, the absence or extreme weakness of interinstitutional accountability leaves electoral accountability as the only instrument for guaranteeing this dimension of quality democracy. The chances of exercising electoral accountability, however, are only periodic, and in some cases citizens must wait several years before the next elections. As O'Donnell (1994) notes, the result is that we obtain a "delegative democracy"—a democracy of poor quality in which the citizen casts his or her vote and is subsequently ignored until the next election. Citizens are left without any means of controlling corruption and bad government, and there are no other institutions really capable of guaranteeing interinstitutional accountability.

Participation

Participation can be subverted and constrained in a variety of subtle and overt ways in democracies around the world. Citizen dissatisfaction, passivity, indifference, and alienation are key reasons for the consequent sharp decline of voting and other forms of citizen participation. But the subversion of the meaning and consequence of participation can be seen when it is no longer spontaneous, voluntary, and free but instead comes to be influenced and even shaped by a different sort of elites. A key role in this can be played by television and other mass media. The so-called audience democracy is the main context of such pseudoparticipation. In fact, within a highly personalized politics where communication elites are very important and the political debate is transferred from institutional arenas to public opinion, effective participation has almost no room. Moreover, there are attempts to secure a controlled participation that may just take the form of obedient support for government actions. That is, there is an effort to get people to participate but only with behaviors that support the incumbent authorities. Other forms of participation are discouraged, and this is not difficult in social and political contexts with a poor tradition of active, autonomous civil society. Participation compounded by various forms of violence is also a subverted way of "taking part" in politics. As suggested by Dahl some years ago (and, more recently, in 1998), a key, necessary, definitional element of democracy is a firm commitment to "the peaceful solution of conflicts." Consequently, the use of violent means twists and distorts the very working of every democracy.

Competition

There are also a few recurrent patterns of subverting competition. The first one is the attempt to exclude competition in some area where the effective working of competition is supposed to have relevant consequences, for example, by making a pact between two parties participating in an election or by agreeing to exclude a priori a political actor, person, or group from fair participation in an election. Second, a distortion of competition can be the end result of inadequate implementation of rules regulating electoral campaigns and financial support for parties. A third, recurrent way of distorting the competition is in obfuscating the program and/or policy differences among parties or party coalitions. Collusive pacts between government and opposition may also be formed. Especially in hybrid regimes, competition can be seriously subverted if leaders and parties are able to arbitrarily control the implementation of rules, especially the electoral ones, or are able to constrain pluralism of information. Finally, use of violence is another way of subverting the competition among political actors.

Freedom, Equality, and Responsiveness

Without going into details, freedom and equality are subverted when they are merely formally acknowledged as rights but not put into practice. The failure to allocate the funding needed for implementation of these rights is one common barrier to their realization. Responsiveness can also be subverted when citizens are not adequately informed about the impact of political actions, as such information is at the root of the formation of responsiveness perceptions.

Empirical research on the varieties of ways the dimensions or qualities can be subverted can also make it possible to detect the democracies *with lesser or without qualities* and even to understand from a different perspective how and why problems of delegitimation and eventually related problems of consolidation can emerge in the scrutinized country.

Leonardo Morlino
LUISS Guido Carli
Rome, Italy

See also Accountability, Electoral; Accountability, Interinstitutional; Competition, Political; Democracy, Theories of; Democratization; Equality; Equality, Political; Human Rights, Comparative Perspectives; Hybrid Regimes; Participation; Responsiveness; Rights; Rule of Law; Transition

Further Readings

Altman, G. A., & Perez-Liñan, A. (2002). Assessing the quality of democracy: Freedom, competitiveness and participation in eighteen Latin American countries. *Democratization, 7,* 85–100.

Beetham, D., Bracking, S., Kearton, I., & Weir, S. (2002). *The IDEA handbook on democracy assessment.* The Hague, Netherlands: IDEA/Kluwer Law International.

Dahl, R. A. (1998). *On democracy*. New Haven, CT: Yale University Press.

Diamond, L., & Morlino, L. (Eds.). (2005). *Assessing the quality of democracy*. Baltimore: Johns Hopkins University Press.

Landman, T. (2006). *Studying human rights*. London: Routledge.

Lijphart, A. P. (1999). *Patterns of democracy: Government forms and performance in thirty-six countries*. New Haven, CT: Yale University Press.

Maravall, J. A. (1997). *Surviving accountability* (Jean Monnet Chair Paper No. RSC 97/46). Florence, Italy: European University Institute.

Morlino, L. (2009). *Qualities of democracy: How to analyze them* (Studies in Public Policies No. 465). Aberdeen, UK: University of Aberdeen.

O'Donnell, G. (1994). Delegative democracy. *Journal of Democracy, 5*(1), 55–69.

Sawer, M., Abjorensen, N., & Larkin, P. (2009). *Australia: The state of democracy*. Annandale, VA: The Federation Press.

Democracy, Theories of

Theories of democracy consist of definitions and generalizations used to describe, explain, and evaluate existing or past political orders. Their purpose is thus twofold: (1) to get a better understanding of the preconditions and performance of democratic systems and (2) to try to judge this reality according to certain democratic core values.

Democracy means collective self-determination. Its purpose is to form political decisions according to the will of the citizens. Apart from this rather abstract meaning, there is a great difference between ancient and modern democracy. The first section of this entry recapitulates the mainly negative use of democracy in ancient thought and the development from the ancient to the modern understanding of democracy, outlining three semantic transformations from the ancient to the modern understanding. The second section focuses on contemporary theories of democracy, distinguishing three types: empirical, positive, and normative. Finally, two different answers to the postnational challenge of democracy are outlined: the concepts of global democracy and postdemocracy.

Ancient and Modern Theories of Democracy

Ancient Critics of Democracy

In ancient political theory, "democracy" was a polemic and negative concept. In the late 5th century BCE, Pseudo-Xenophon, the first Athenian critic of democracy we know of, calls it a regime in which the many rule in a selfish and destructive manner. Democratic men strive for their personal gain, and to this end, they not only suppress the aristocratic best and the population of the naval colonies but also rule without regard to the common weal. According to Pseudo-Xenophon, this leads to an unjust political and moral order. However, he does not deny a certain rationality of the democratic regime insofar as the many live better in a democracy. Thus, he ends with the paradox that from an aristocratic perspective the democratic regime is clearly unjust, whereas the same order seems at least internally rational from a democratic point of view.

It is in Plato's work that this paradox is resolved inasmuch as he develops a metaphysical foundation of the political order. In concurrence with

Pseudo-Xenophon, he describes the supposed shortcomings of the democratic regime, such as rhetorical betrayal and demagoguery. Because the many cannot know the political areté, they are merely a pawn for the sophists who are interested only in their personal gain. To counter these democratic practices, Plato formulates his model of the philosopher-king, which is founded in an epistemic understanding of politics. For Plato, good politics is based on higher knowledge, which only the few with special philosophical talent and training can attain.

There are at least two problems with this fusion of politics and philosophy. The first is its metaphysical character. Plato bases his political philosophy in an idealistic framework dubious even to his contemporaries. The second problem is the utopian character of the philosopher-king, as Plato himself mentioned in the *Politeia*. Democratic practices and norms are very widespread in the Athenian demos, so his model is simply not realizable. Therefore, Plato argues in the *Laws* for a second-best regime. It is no longer the philosopher-king who is to guarantee the good order but a system of laws that regulate even the smallest details. To integrate the many into the regime, he considers democratic modes of decision making, such as the participation of the many in elections and even drawing lots. However, these concessions to democratic practices do not mean that Plato has abandoned his aristocratic ideals. Democratic institutions are subordinated to exclusive ones, thereby thwarting the rule of the many.

Aristotle's reflections on democracy differ from Plato's in at least two ways. First, he develops his insights by examining the empirical world. Second, he transcends Plato's dichotomies of philosophy and democracy, the few and the many, and formulates a more integrated and therefore realistic understanding of the political world. Aristotle even concedes that the democratic order displays some rationality insofar as the many, if they deliberate together, can obtain more information than an oligarchic assembly can. Nevertheless, he is no friend of democracy. In *Politics*, he criticizes the democratic order of Athens, which he regards as "extreme," and as Pseudo-Xenophon does, he criticizes demagoguery and the tendency of the demos to neglect the common weal. Therefore, Aristotle argues for a more moderate constitution, the "polity," which he sometimes also calls the "best" democracy. This constitution is characterized by the rule of law, and oligarchic and democratic institutions are mixed. In this regime, a strong separation exists between the political experts—that is, the educated few, who are chosen mainly by elections and not by lot, and the orderly demos. With this conception, Aristotle abandons the Athenian praxis of direct democracy and moves in the direction of a modern, representative understanding of democracy.

In short, for these "intellectual critics of popular role" (Josiah Ober, 1998), the democratic praxis of the ancient world with the direct involvement of the demos was unjust and highly pathological. This assessment did not change after the decline of the Grecian city-states at the end of the 4th century BCE. On the contrary, this change further supported the antidemocratic bias of political thought. The reference point was no longer the political praxis of the city-states but the antidemocratic writings of Thucydides, Plato, and Aristotle. In the Roman discussion and the political theory of the Middle Ages, democracy was not only regarded as a regime of the past but also as an illegitimate order. In the early-modern age, this critical or even negative assessment of democracy changes only slightly. Until the end of the 18th century, the pejorative connotation of "democracy" dominated political thought. However, after the "democratic" revolutions of the late 18th century, there was a new beginning in democratic theory. It was a beginning that fundamentally changed the semantics of democracy from the direct involvement of the masses to a system in which elected representatives rule.

This shift in meaning from the antique to the modern concept of democracy was carried out in a multistage process of transformation. The basic semantic changes concern the evaluation, temporalization, and institutionalization of the concept of democracy and can be described with the terms *positivation*, *futurization*, and *completion*.

Three Semantic Transformations of "Democracy"

Positivation

In the ancient theories of Plato and Aristotle as well as Cicero and Polybios, "democracy" was a negative concept. All major primary sources from which the ancient concept of democracy is handed down to us are by critics, if not enemies, of democracy. Their critique was vehement, and their list of democracy's shortcomings contained, as shown, very different points: It permits unqualified citizens to participate in politics, it complicates political decision-making processes, it produces bad decisions, it debauches the political culture, or it is simply an amoral order—just to mention the most important points of criticism. This negative usage of the concept continued uninterrupted from the Middle Ages to modern times, and only in the writings of Spinoza and in the political speeches of some Dutch Republican thinkers in the 1780s can one find attempts to give democracy a positive designation. This positive connotation of democracy gradually became accepted after the French Revolution and then in the course of the extension of suffrage in the United States, Western Europe, New Zealand, and Australia in the 19th century. This process was accompanied by ideological disputes that ended in the mid-20th century. Today, the transformation to a positive concept is complete at least in Western society; the concept has developed into a category of self-description in global political disputes. While democracy in modern democracies has many critics, it no longer has any fundamental enemies, at least in the Western world.

Futurization

Even the political thinkers of Hellenism and later Roman authors such as Cicero regarded democracy as a form of government of a bygone era. They considered it a thing of the past and associated it with the existence of small city-states of the lost world of ancient Greece. If only for that reason, and regardless of its negative aspects, authors such as Baron de Montesquieu, John Locke, or the writers of the Federalist Papers did not think it a serious option as a concept for the political future. Admittedly, Montesquieu arrives at a positive understanding of democracy in his *Spirits of the Laws*, but he binds the idea of democracy so tightly to the prerequisites of equality and rurality that there is no room for a democratic order in his age with its advancing economical and social differentiation. For him, democratic practices could only be a subordinate element in a mixed regime. Montesquieu's understanding is typical for the equation of democracy with the praxis of the ancient city-states. For him, democracy was a regime of the past and not a realistic option for the future. Georg Wilhelm Friedrich Hegel had a more positive view of the contemporary meaning of ancient democracy, but since the liberation of the subjective mind, he too was unsure of its future. Even authors such as Johannes Althusius and Jean-Jacques Rousseau, whose theories ventured to take big steps toward positivation, were rather cautious concerning a realistic future for democracy.

It was only with Alexis de Tocqueville's book on America that a political rhetoric prevailed that turned around the structure of time dominating in most contemporaries' minds and enabled them to see democracy as a project of the future. In Tocqueville's view, North America was already mostly a democracy, and Europe would soon be predominantly a democracy as well, as problematic as he felt this tendency to be. This futurization made the concept of democracy a key term for the political battles of the 19th and 20th centuries. This perspective electrified enemies as well as proponents of democracy, the former because they now faced a challenge that laid claim to the future, the latter because they had a feasible political project with the name "democracy" before them. Today, the futuristic character of democracy is undisputed. Democracy is a project on perfecting in which we all cooperate, in the hope of one day accomplishing it completely.

Completion

Third, the concept of democracy underwent a fundamental change in its institutional inventory. While there was a primacy of political participation in antiquity, slowly a constitutional usage prevailed that systematically restricted the moment of direct participation. It is the transition, so welcomed by Benjamin Constant, from the freedom of the old to the freedom of the new at the beginning of the 19th century that makes this paradigmatic rupture apparent. The change from a negative to a positive evaluation of the concept of democracy coincides historically with the transition to a primacy of liberal defensive rights and the installation of a representative system. Democracy is now regarded as an institutional order that must be complemented by a system of "checks and balances" so that negative freedom—the protection of the individual from decisions by the democratic majority—is secured. Accordingly, the list of proposals of how the institutions of democracy should be complemented is long and bears witness to a high level of institutional creativity on the part of contemporary authors. The most important ones are constitutionalism (e.g., independence of the judiciary, a coherent legal system), different models for the separation of powers, federalism, and multistage representative systems.

It is only because of these three semantic transformations that a concept of democracy, which stands in such conspicuous discrepancy to its original usage in antiquity, could survive. In view of the great shifts in meaning, it is hard to answer the obvious question of why the concept of "democracy" was not simply given up, instead of being intricately filled with new meanings. Its astounding ability to survive can probably be explained best by the attraction of associations inspired by the parts of the Greek compound—"demos" and "kratein." The rhetorical reference to "the people" and their "rule" constitutes a—however weakened—reference to participatory components in political systems and provides them with mass legitimacy.

Modern Theories of Democracy

Empirical, Positive, and Normative Theories of Democracy

There are different ways to comprehend the extensive and confusing debates of modern theories of democracy. Basically, the discussions can be grouped in two approaches: diachronic and synchronic ones. The first approach traces the historical development of democratic thought. The purpose of this method is to detect the crucial steps and the striking changes of democratic thought. The "advancements" of modern theories of democracy are shown, for example, the history of the concept of representative democracy from John Locke to Robert A. Dahl. In the second approach, theories of democracy are condensed to models or paradigms, for example, liberal or republican theories. The purpose of this method is to compare the different models and to rank them. This ranking can take place according to their degree of accordance with the institutions of liberal democracy or the requirements of global governance. However, there is another distinction underlying these differences. It concerns the scientific modality of the theories, the way they look at democracy. Therefore, one can distinguish between three "logical" modes of democratic theory: the empirical, the positive, and the normative type.

Empirical Theories of Democracy

Empirical theories of democracy try to rank political systems according to a scale of democratic values and institutions or to determine the necessary functional preconditions of democratic systems and measure how such systems perform.

The goal of the first group is to construct reliable and standardized scales in order to obtain a yardstick for comparing different political systems that can then be ranked according to their degree of democracy. Dahl wrote the classical study in 1971. Dahl formulates seven indicators of a "polyarchy," which in his view is the modern form of democracy: (1) the freedom to form and join organizations, (2) freedom of expression, (3) the right to vote, (4) the right of political leaders to compete for support, (5) alternative sources of information, (6) free and fair elections, and (7) institutions for making government policies responsive. The points of reference for these democratic standards are, on the one hand, the norms and institutions of Western democracies and, on the other, normative theories of democracy, which try to justify these norms and institutions. In the next step, Dahl measures political systems according to these indicators and then orders them, applying a scale ranging from full polyarchies to near-polyarchies to nonpolyarchical systems. Dahl's empirical finding is that the Western democracies are not the only full polyarchies but also countries such as India and Costa Rica.

Dahl's work inspired many empirical studies based on different theoretical foundations. Especially in recent years, measuring democracy has become a burgeoning academic pursuit. The discussion is this field is focused on adequate indicators of democracy and their application in empirical research. At this point, a fundamental problem of measuring democracy arises: The selected indicators as well as their operationalization rest on more or less carefully considered normative assumptions about the nature of democracy (see below).

A second group of empirical theories tries to analyze the functional preconditions and the performance of democratic systems. Their starting point is the sociological theory of modernization, which in the 1950s asserted that the connection between economic development and the political system of societies was narrow. In 1960, Seymour Martin Lipset wrote one of the first studies in this field. He investigated the connection between the degree of democratization and social-economic variables such as average income or level of education. His main result confirmed the theory of modernization: The higher the level of economic development, the more democratic a society will be. Many scholars have followed Lipset and, with some variations in research design, have come to similar conclusions. Compared with autocratic systems, democratic ones perform much better in most policy areas (e.g., health care, education, or environmental resource management). Such research can be carried out contrasting democracies with other political orders such as monarchies or one-party systems. Another way is to estimate the performance of different subtypes of democratic systems, for example, parliamentarian or presidential systems.

Positive Theories of Democracy

Positive theories of democracy construct formal models of the democratic process, for example, voting behavior. They do not involve empirical study of the workings of real democracies but rather are deductive theories of political processes under constructed conditions, such as the rationality of agents or the closed logic of functional systems. Their starting points are certain axioms that are used as a basis for developing the main characteristics of democratic systems. In contrast to empirical theories, neither the basic assumptions nor the causal or functional explanations of positive theories claim to be normative. Rather, the authors of positive theories explicitly do not want to formulate normative statements; they seek only to characterize the democratic process and to explain typical political sequences in modern democracies.

There are two versions of positive democratic theories, and they focus on the opposite ends of the democratic process: One is based on the findings of rational choice theories; the other is Niklas Luhmann's systems theory. Rational choice theories focus on the microlevel of society and start with the assumption that individual actors are mainly motivated to maximize their personal gains. With the same theoretical grounding, they try to explain the political actions of collective actors such as political parties, interest groups, and even states. In contrast, systems theory begins on the macrolevel of society. Luhmann regards the different realms of society, for example, the economy, the political sphere, or the scientific community, as distinct and self-contained systems and analyzes their structures as well as their functions for the whole of society. Individual actions are outside of the theory's focus, they are simply irrelevant. In Luhmann's view, systems function according to their own logic, which is independent of the actions of individual persons.

Even though rational choice theory and systems theory try to avoid normative claims, they nevertheless possess enormous critical potential because they demonstrate two shortcomings of democracy: the irrationality of democratic decision making and the constrained range of political actions. One of the most important normative arguments for democratic modes of decision making is the assumption that one can determine the correct will of the majority in this way. However, according to the findings of Moisei Ostrogorski, one of the forerunners of rational choice theory, this is a myth. The Ostrogorski paradox shows that even the smallest changes in voting behavior can lead to big differences in voting results. This finding is explosive from a normative point of view because it casts doubt on the legitimacy of majority decisions. Besides that, rational choice theory does not regard irrational decision as deviations from democratic norms but as the inevitable result of the aggregation of individual votes or of the merging of different forms of rationality in the political process.

Similar statements can be made about the findings of systems theory. Although Luhmann regards democratic political systems as the appropriate form of modern, functionally differentiated societies, he criticizes the normative bias of democratic practices. In his view, the political system in modern societies is no longer at the top of a pyramid but is just one system among others, with a code of its own. In addition, this code of government and opposition cannot be transferred into other functional systems, for example, the economic or the juridical system, without harm. Therefore, the effects of democratic politics are limited. According to Luhmann, democratic imperatives simply cannot govern the other functional spheres, the code of government and opposition can only irritate, at worst destroy, the reproduction of the other systems. The function of the political system and its democratic form is thus only to reproduce the necessary illusion that a society can be governed by collective decisions.

In contrast to empirical and positive theories, normative approaches try to formulate convincing justifications of democratic orders. The goal of normative approaches is to deliver criteria for praising or criticizing normative and institutional orders. In this way, empirical theories can use their findings to evaluate existing political systems. Normative theories explicitly do not strive for ethical neutrality or freedom from value judgment. Far from it, these value-based justifications are seen as an essentially scientific endeavor. Some scholars cast doubts on the scientific character of normative approaches due to this ethical grounding. However, as seen, even positive and empirical theories are not value-free, and it can be said that normative justifications elucidate the inescapable nature of any type of democratic theory.

The Three Normative Axes of Theories of Democracy

Normative theories of democracy differ from each other, too, in regard to the way they are reasoned as well as in their institutional consequences. In the past 20 years, we have witnessed a tremendous differentiation and refinement of the traditional models. In addition to the classical approaches of liberal, elitist, conservative, socialist, and participatory theories of democracy, deliberative, neorepublican, neoliberal, communitarian, cosmopolitan, associative, feminist, ecological, experimental, multiculturalist, and postmodern theories have come along, to mention only the most important ones. The subjects of the normative theories are by now all imaginable aspects of democracy: its traditions, goals, institutional settings, and procedures. Normative theories have been and still are largely influenced by political fashions whose themes are at the center of scientific debates. Therefore, a listing and sorting of these debates offers only little orientation in the jungle of the numerous normative theories.

Rather than simply listing them, it is more helpful to ask how normative theories proceed in assessing real or hypothetical political arrangements as "democratic." In other words, how do they construct the criteria that are to generate normative statements about political systems and procedures? Obviously, the development of theses criteria does not take place in a vacuum but is incorporated in certain political experiences and assumptions about the central problems of democracy. Thus, they have historical underpinnings. The forms of those experiences and assumptions that have structured the debates in the past 3 decades can be named the "three normative axes of theories of democracy."

The Social Object of Democracy

On the first axis, the discussions concern the social spheres in which democratic norms should and could apply. The recent debates about the appropriate social object of democracy focus on four issues. Common to them is an inclusive perspective—that is, the integration of additional social spheres under democratic rule. An initial focus in the 1970s was the call to democratize further realms of society, especially the economy, the workplace, and educational organizations. These demands found a certain resonance in theories of "strong" or "radical" democracy. In the 1980s and 1990s, feminist theories of democracy were a second focal point. Starting from the empirically based assumption of a "gender gap" in modern democracies, their goal is to include females and female perspectives in the political process. In most feminist approaches, the fulfillment of this demand is connected with a qualitative advancement of politics and policies. These positive expectations are related to a new, more communicative style of politics and a political agenda directed to making family life and the working environment more compatible in modern societies. In a third debate about inclusion, and in parallel to the feminist theories of democracy, various concepts of multicultural democracy have been developed. Based on the observation that minorities in liberal democracies are discriminated in many ways and that they are not adequately represented, the purpose of multiculturalist approaches is to integrate the different collective identities in the political sphere. The proposals for reform are centered on the opening of existing institutional arrangements. Some authors propose special mechanisms of group representation or quotas that have to be balanced with the liberal-democratic demand for equality.

In a fourth and more recent discussion, some authors argue for an inclusion of children, future generations, and even apes in the democratic process. At the very least, the demand to consider the interests of apes in political decision making, which seems odd at first glance, highlights the difficulties of obtaining adequate criteria for limiting the social object of democracy. In the semantics of "democratization," democracy is in principle a never-ending process during which its boundaries and goals must be discussed again and again.

Degree and Ways of Participation

On the second axis, normative theories of democracy formulate statements about the optimal degree of participation and the ideal relationship between the entire demos and the political elites. The focus of all relevant controversies about democracy during the 20th century was the degree and the forms of citizens' involvement in the political process: the debates between the advocates of government by council democracy and those of parliamentarism in the 1920s and again in the1970s, the political disputes between the champions of a Fascist leader and their Liberal and Leftist adversaries in the 1930s, the controversy between the advocates of a representative and those of a plebiscitary democracy in the 1960s and in the 1980s, or the dispute between the theorists of an elite democracy and those of a grassroots democracy in the late 1960s and early 1970s and again in the 1990s. In all those debates, it was and still is contested how the demos is to be included in the process of political decision making. Authoritarian concepts consider the process of acclamation to the charismatic leader to be the genuine democratic way of declaring consent; liberal and elitist conceptions link their versions of democracy with the principle of representation and elections; and theorists of a grassroots democracy or of a government by council democracy support the idea of an intensely participating citizen.

While the controversy about the pros and cons of a higher degree of citizen participation in the 1970s and 1980s were hot-tempered and ideologically based, it has cooled down since then. One of the main reasons for this development is the findings of empirical research on direct democracy. Neither the hopes of the advocates nor the fears of the adversaries have been fulfilled. The empirical results show instead that procedures of direct democracy can lead to a higher degree of citizen satisfaction and rational problem solving in certain cultural and institutional contexts, but they can also be used as an instrument of populism. Because of these complex and contradictory findings of empirical research, the question about the forms of citizen involvement is no longer at the center of normative theories of democracy. By now, the mainstream literature focuses on the institutions of democracy and the question of their performance. This leads us to the third axis.

Degree of Rationality

On a third normative axis, theories of democracy finally contain certain assumptions about the degree of rationality of democratic decision making. They make statements about the technical and factual reasonableness of democratic decisions and even about their moral quality. Thus, conservative and liberal critics have frequently accused democracy of leading to irrational decisions. Leftist advocates of democracy, on the other hand, have considered the democratic character to be the ultimate condition that makes correct and therefore rational decisions possible. Only in recent discussions have these factional struggles faded away.

The rise of the theory of deliberative democracy in the past 2 decades has demonstrated that the question of rationality has gained relevance significantly. Deliberation is the exhaustive and reflective debate about political questions. The deliberative give-and-take of arguments aims to elucidate individual and collective interests. More ambitious forms of deliberative democracy do not stop at this point and demand of the participants that they transcend their own wishes in view of new moral insights. This change of the guiding interests is attributed to the public nature of deliberative processes and is characterized as a clarification and moral bettering of preferences. Although the goal is not to discover an unchangeable moral truth, the expectation is that in the deliberative process all arguments that serve only private goals can be eliminated. For this reason, advocates of deliberative democracy claim that its results have a higher degree of legitimacy than elections and voting.

Theories of deliberative democracy stand against two alternative theories. On the one hand, they criticize rational choice conceptions of the aggregation of preferences and insist that preferences can change in communicative processes. On the other hand, deliberative approaches fault models of democracy that demand a higher degree of participation without showing how the citizen can gain the moral resources required for this ambitious endeavor.

The institutional implications of deliberative approaches go in three directions. The first is directed toward the individual citizen. According to Robert E. Goodin, each member of the political community should reflect on the moral implications of his or her preferences ("deliberation within"). The practical consequence of this approach is the demand that all citizens receive a better education, which is to lead to higher sensibility for the moral interests of the other citizens. A second strain argues in connection with Jürgen Habermas for the deliberative character of political institutions as representative assemblies and channels of the public sphere. These approaches do not strive for a basic change of the political institutions but for a widening of the deliberative character of liberal democracies, for example, in strengthening the discursive character of parliaments, or for enhancing the rationality of public debates. Other authors argue for institutional innovations within liberal democracy. One prominent suggestion is the deliberative opinion poll by James Fishkin, where randomly selected citizens discuss political questions in a deliberative setting. The results of these deliberations can inform the elected authorities or even lead directly to political decisions.

A third and more radical version is not satisfied with such a deliberative interpretation of existing societies but argues for changes in the basic structures of liberal democracies to achieve the moral goals of deliberative democracy. They criticize the different forms of exclusion in modern societies that are still relevant in deliberative settings. Therefore, they demand a democratization of all of society as a prerequisite for deliberative procedures. These radical approaches have some obvious connections with theories of participation and grassroots theories of democracy, and they support a politicization of civil society.

Interplay of Empirical and Normative Theories and the Rationalization of Democracy

Empirical and normative theories of democracy are of course not entirely separated. As seen, empirical theories draw their criteria for assessing the democratic nature of political systems from theories about the essential norms and institutions of democracy, such as Dahl's. Empirical democracy research is assisted by systems theory and rational choice theory, each of which in its own way theoretically deduces why political participation in modern mass democracies is nearly without effect and ultimately makes no sense at all. Nevertheless, this is only one side of their influencing one another; contemporary normative theories in turn are oriented to the findings of empirical research on theories of democracy.

All normative approaches that can currently claim scientific relevance follow the findings of empirical democratic theory. In the academic debate about democracy, no idea is taken seriously if it does not demonstrate its ability to be connected to empiricism and thus its proximity to reality. Thus, the findings of empirical research on democracy, for example, on the degree of political interests or on the irrationality of most citizens' political preferences, produce subliminal but nonetheless powerful pressure on the making of normative theories. In this perspective, democracy becomes a regime type that produces a certain amount of legal certainty, cultural and educational goods, welfare, and other collective goods (recently, above all, security from terrorism) but that has lost the active political participation of its citizens.

With this result, normative democracy theory again provides the justification for empirical democracy research in which rule of law and the production of welfare and stability constitute the most important parameters, more important than the participation of all citizens: At best, participation is used as a dependent variable. The indices of empirical democracy measurement stubbornly test for certain institutionalizations, for example, the existence of individual rights or of the basic building blocks of a parliamentary democracy with separation of powers. The participatory component has become a ballast of the concept of democracy, standing in the way of its continued success. When reviewing the most important approaches in current political theory, one notices—despite all the differences—a common terminological shift with which the path to a fourth and new semantic transformation in democratic theory was paved. Most current theories use a concept of democracy that discerns deep chasms between political participation and "rational" results and, when in doubt, argues against political participation.

This transformation can be described as the "rationalization" of democratic theory. It means that the focus of modern theories of democracy has shifted to the evaluation of the quality of the results of politics. Democratic theory is becoming more and more output oriented, and its normative efforts' main goal is to increase the degree of rationality of this output. Larger differences within this paradigm occur only where the following criteria of rationality are concerned: effectiveness, feasibility, representation of interests, justice, or the public weal. Political participation is no longer regarded as the goal but as one of several possible ways to enhance the degree of rationality of collectively binding decisions.

The Future of Democratic Theory: Global Democracy or Postdemocracy?

One of the most important motivations to "rationalize" democracy is the decline of the nation-state in the era of globalization. In the past 2 decades, we have witnessed a displacement of political decisions from the national to the international or supranational level. This is of course a reaction to the growing need for global coordination of economic and political processes. The answer to this postnational constellation is the development and strengthening of regimes and organizations as means for "global governance." The most impressive of these supranational organizations is the European Union (EU) with a huge amount of political competencies. However, even at the global level, there is a large variety of regimes and organizations for coordinating international politics. One only has to think about the World Bank, the International Monetary Fond, and the World Trade Organization to comprehend the relevance of international organizations for the well-being of nearly every person on the planet. However, these global economic multilaterals are only the best known organizations of global governance. It seems that the era of the nation-state has gone; at least the political power of the states to determine the lives of their citizens has been weakened drastically in the past 20 years. This raises the question as to whether postnational forms of democracy are imaginable.

Theories of Global Democracy

The process of globalization is a major challenge for democratic theory because most of the conceptions discussed above—with the exception of deliberative democracy, see below—have taken the national base of democracy for granted. In addition, the all-important institutions of modern democracy such as elections, parliaments, political parties, and the public sphere are embedded in the nation-state. Nevertheless, the era of the democratic nation-state seems to be coming to an end. Accordingly, there is a growing literature on the question of the democratization of global governance. One of the issues discussed most is how to transform the democratic institutions and procedures to the supranational and even the global level. At least three strains of the debate about democracy in the global order can be discerned.

The first group of authors doubts that the new structures of global governance are responsible for the significant decline of democracy we are witnessing today. In contrast, the second group begins by asserting a fundamental democratic deficit and tries to transfer the values and institutions of national democracy to the supranational and global sphere. In addition, a third group makes the effort to change the semantics of democracy in order to demonstrate the perspectives global governance may open up for democracy. In the context of the debate about a so-called democratic deficit of the EU, but also in view of the institutions of world politics, authors such as Giandomenico Majone and Andrew Moravcsik argue that this accusation is misleading for two reasons. First, there exists a kind of legitimating chain from the direct or indirect election of national parliaments and governments up to the supranational institutions such as the European Commission. It is a kind of political delegation of power from the constituents in each country to their national representatives and delegates in the transnational or global settings. Therefore, even the World Trade Organization is considered democratically responsive. In a second and supporting way, Majone and Moravcsik argue that many of the supranational and transnational institutions of world politics are directed toward technical issues that should be depoliticized. Examples include the world finance institutions or the European Central Bank or institutions occupied with juridical and economic issues. In these authors' view, it is in the interest of the people that experts manage these issues, because only they have the necessary knowledge. It is a governing not by but for the people. Countering this opinion, many scholars have objected that they have overlooked the fact that juridical and economic issues of global governance are inevitably political. Moreover, without the involvement of the people, there is the risk that these bodies of experts will consider only the interests of strong actors.

Thus, in contrast to Majone and Moravcsik, a second group of authors such as David Held and Daniele Archibugi begins by claiming that there is a real democratic deficit in global governance. Their theoretical background is the norms and institutions of national democracy, and they try to ascertain how these norms and institutions can be transformed to the supranational and even the global sphere. Only if the global order can be subjected to democratic values, such as the equality of citizens, the majority principle, and the duty of governments to act in the interests of the people, can the new forms of governance be legitimate. The crucial question in the global age therefore is, "How can democracy preserve its core values and yet adapt to new circumstances and issues?" (Archibugi, 2004, p. 446). The answer to this question is the concept of cosmopolitan democracy, which is conceived as a multilevel order. There should be different degrees of democratic participation at the local, national, interstate, supranational, and global levels. In the concept of a cosmopolitan democracy, the participation of the people at the local and the national levels is to follow the traditional understanding of democratic norms, institutions, and practices, but this is also demanded regarding the supranational and the global levels. Advocates of cosmopolitan democracy argue that there is an emerging global public sphere, consisting of international nongovernmental organizations (INGOs) such as Greenpeace and Amnesty International, as well as a global media system. Even if these structures do not fulfill the demands of a democratic public sphere, they are at least the necessary conditions for strengthening and increasing the transparency and accountability of global politics. However, these are only the preconditions of a cosmopolitan democracy. Both at the supranational and the global levels, parliamentary institutions are regarded as the cornerstone of a democratic order. For the advocates of a cosmopolitan democracy, a world parliament is thus the sine qua non. This parliament is to be elected by all citizens of the world (at least those of democratic states) and have the right to make legislation valid worldwide. One of the logical necessities of this model is therefore a new understanding of the role of states. They are no longer sovereign actors but only one vehicle of the democratic governance of the people. Some authors even foster the notion of a world-state, and therefore some kind of world government, to implement global justice and democratic demands. Opponents of cosmopolitan democracy fundamentally doubt the prospects of a democratic world order. They refer mainly to the structure of power in world politics and to the enormous obstacles to the participation of the people in global governance. Nearly 6 billion voters, for example, would elect a world parliament. How can one guarantee the representative nature of this parliament? In addition, what about the populations of nondemocratic systems—today nearly three quarters of the world's population? The critics of cosmopolitan democracy view even the global public sphere and INGOs critically. For others, the project of cosmopolitan democracy is just a new way to establish and strengthen the hegemonic power of the West and the values and practices of a new global class as cosmopolitanism of the few.

A third group also begins by claiming a democratic deficit in global governance, but, in contrast to the advocates of cosmopolitan democracy, they do not want to transfer the institutions of national democracy to the global level. Instead, these authors try to reconsider the notion of democracy in order to close the legitimacy gap in world politics. There are at least two ways of redescribing democracy along these lines. The first consists in highlighting new ways of political engagement. According to authors such as Ann-Marie Slaughter, international non-INGOs are the representatives of a global demos. Together with the postulate of a growing global public sphere and a growing process of juridical constitutionalization, INGOs are seen as an essential component of a new democratic world order. Their inclusion in the processes of global governance, for example, in the hearings of the World Trade Organization, is seen as a crucial step toward the democratization of international politics. Critics have countered this scenario by arguing that a global public sphere and a vivid array of international civil actors are only the precondition but not the essence of democracy. Others criticize the oligarchic structure of most of the INGOs and that they mostly articulate the demands of the rich countries of the Western world. The question therefore is, Who has authorized these organizations? Furthermore, critics object that the influence of the INGOs in international politics should not be overestimated. According to these critics, world politics is a game played only by powerful elites and their experts. In their view, the participation of nongovernmental organizations merely functions to legitimize undemocratic ways of decision making.

A second way of adapting the meaning of democracy to the needs of global governance is taken by the advocates of a deliberative understanding of politics. Inspired in particular by the work of Habermas, some authors rely on the epistemic functions of transnational decision-making bodies. The starting point is Habermas's distinction between a substantial and a proceduralist understanding of popular sovereignty. Habermas refers to an anonymous civil society without a concrete democratic subject. This opens up the possibility of transferring the concept of democracy to the global sphere where the demos is only conceivable in form of the manifold demands articulated there. The crucial step is the integration of these demands into the global decision-making bodies. In addition, a second feature of deliberative democracy is relevant here—namely, the epistemic understanding of politics. As seen above, the goal of the deliberative process is to tease out the better argument by means of discussions between the relevant groups. But it is often difficult to include all relevant groups in international politics. However, according to the advocates of deliberative democracy, this need not lead to an undemocratic way of decision making. As far as the relevant viewpoints are included, it is sufficient that political experts discuss the relevant topics to gain "rationally acceptable results." Critics argue against this understanding of democracy, insisting on the necessity of real participation by the people. Habermas's legitimation of the epistemic gains of global governance threatens to erode the very meaning of democracy—that is, the involvement of all the people and not only the advocates of the better argument.

Postdemocracy and Beyond

Another way to capture the changing reality of democracy in the era of globalization is the strategy of Colin Crouch, who describes Western political systems as "postdemocracies." According to Crouch, genuinely democratic institutions such as parliament, regular elections, party competition, and the rule of law still exist. Therefore, these societies differ in a significant way from autocratic societies. Nevertheless, the processes of globalization and the weakening of the state's capacity to regulate the economy are progressively undermining these institutions. This leads not only to a loss of importance of central democratic ways of political decision making but also to a shift in power relations. According to Crouch, Western societies are therefore characterized by a fundamental ambivalence. On the one hand, the forms of democracy still exist, and are even expanding (e.g., due to the establishment of new forms of direct democracy), but on the other, they have lost their relevance and are getting wedged in by new intransparent forms of national and global governance, where powerful elites dictate the rule of the game. The most significant consequences of this new power structure are the decline of egalitarian politics and the expansion of new forms of political marketing.

Thus, for Crouch the use of the term *democracy* only obfuscates the fact that in Western societies, too, the vital circumstances that influence the lives of citizens can be decided collectively only to a very small extent—a fact that was conceded in the professional discourse of political science long ago. This generates not only expectations but more or less subliminally also a deceptive appearance.

The thesis of postdemocracy is a polemical reaction resulting from the desire to draw attention to the fact that modern Western political systems have drifted away from the basic democratic impulse. However, there are few indications that this strategy can prevail beyond a tight circle of scholars. Rather, one can expect that the promise of democracy will not lose its political force, even in a world in which the national base of democratic practices is weakened.

Hubertus Buchstein and Dirk Jörke
University of Greifswald
Greifswald, Germany

See also Deliberative Policy Making; Democracy, Quality; Equality, Political; Modernization Theory; Normative Political Theory; Participation; Political Philosophy; Rational Choice; Social Democracy; Systems Theory

Further Readings

Archibugi, D. (2004). Cosmopolitan democracy and its critics: A review. *European Journal of International Relations, 10,* 437–473.

Bohman, J., & Rehg, W. (Eds.). (1997). *Deliberative democracy: Essays on reason and politics.* Cambridge: MIT Press.

Crouch, C. (2004). *Post-democracy*. Cambridge, UK: Polity.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Dahl, R. A. (1989). *Democracy and its critics*. New Haven, CT: Yale University Press.

Dunn, J. (2005). *Setting the people free: The story of democracy*. London: Atlantic Books.

Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy*. Cambridge: MIT Press.

Held, D. (1995). *Democracy and the global order*. Cambridge, UK: Polity Press.

Held, D. (1996). *Models of democracy* (2nd ed.). Palo Alto, CA: Stanford University Press.

Keane, J. (2010). *The life and death of democracy*. New York: W. W. Norton.

Lipset, S. M. (1960). *Political man*. New Haven, CT: Yale University Press.

Manin, B. (1997). *The principles of representative government*. Cambridge, UK: Cambridge University Press.

Ober, J. (1998). *Political dissent in democratic Athens: Intellectual critics of popular role*. Princeton, NJ: Princeton University Press.

Rosanvallon, P. (2008). *Counter-democracy: Politics in an age of distrust*. Cambridge, UK: Cambridge University Press.

Young, I. M. (2000). *Inclusion and democracy*. Oxford, UK: Oxford University Press.

Democracy, Types of

Since the end of the Cold War, democracy has become the unrivaled form of government in the world. Acceptance of a country as a full partner in the global community of nations is considerably facilitated by its being characterized as a political democracy; international military interventions, as in the case of Iraq and Afghanistan, state their goal as the building of democracy; less democratic countries are asked to improve the quality of their democracy to gain esteem; and countries that hardly possess the attributes of democratic governance claim to be democratic because such characterization is thought to bestow prestige on them.

Among democracies, there is no single mode of organizing a polity as a political democracy. Institutional arrangements of democratic governance have varied across time and countries. Furthermore, democratic systems have evolved and operated in countries that have, among others, different histories, cultures, traditions, economies, demographic compositions, and socioeconomic characteristics. These factors have all put their imprint on how democratic institutions are organized and operate in specific countries. Attempts at presenting and discussing types and typologies of democracies are several. Some types and typologies, not widely employed by students of democracy thus far, have not been included in this entry. But before presenting the most well-known analyses, a methodological "warning" is necessary.

The word *type* may be used in two different senses. First, it may be used to define the necessary set of characteristics that need to be present to identify certain phenomena as distinct from others and give it a name. The word *model* is also sometimes employed to convey the meaning of type in this sense. Second, it may be a category that emerges from classifying phenomena according to certain criteria. In this second sense, "type" emerges as a subcategory—that is, as the product of an effort at generating a typology. One may, for example, identify all political systems that satisfy the criteria for being a democracy (model) but then group them according to some key features (variables) along which they may differ from each other. This exercise, which aims at developing a typology, generates subcategories that are also called "types." Types as subcategories of a typology include all attributes of democracy and then some that make them unique, helping us distinguish them from others.

This entry is a discussion of the types of democracy. The first section focuses on the two fundamental models of democracy, direct and liberal representative, as major concerns of political thought. This section on normative models introduces the fundamental philosophical underpinnings of actual democratic systems as they have developed in different societies. The second section offers a typology of democratic systems as they operate in the world today. Here the scheme of classification is based on whether minimum or maximum possible majorities are sought in making decisions since such a distinction constitutes a major variable along which contemporary democracies differ from each other. Some types based on the empirical study of how systems function are also offered here. Finally, normative models critical of the outcome of the operation of contemporary democracies are taken up.

Normative and Empirical Models

Two different approaches may be used in defining types and developing typologies. The first approach confines itself to positing a set of criteria or norms, thereby defining what a democracy is or ought to be. Using this normative theoretical approach, one may identify models of democracy. The second approach focuses on societies that already have systems that satisfy the norms that are depicted in the definition of democracy but then proceed to classify them, thereby developing typologies and then types as subcategories based on how each differs from the others, for example, in terms of its institutional arrangements, its cultural and behavioral environment, and its functioning. Understandably, this second approach requires an empirical examination of democratic societies. This approach, in contrast to normative democratic theory, constitutes the basis of empirical democratic theory.

Classic Normative Models of Democracy

Dictionaries, relying on the Greek origins of the word, *demos* (people) and *kratos* (rule), often define democracy as rule by and for the people. More scholarly definitions have elaborated the basic idea. Charles Tilly (2007), for example, has emphasized "the degree that political relations between the state and its citizens feature broad, equal, protected and mutually binding consultation" (pp. 13–14). Such definitions have focused neither on who the people are nor on how or through what mechanisms they rule themselves. The question of who the people are, after historical and long struggles, has been answered by the realization of universal suffrage. How the people should rule themselves, on the other hand, has been given two different answers, leading to the emergence of two normative models of (1) *direct or participatory democracy* and (2) *liberal representative democracy*.

Direct Democracy and Its Variants

In terms of historical order, direct democracy comes first. Its basic idea was developed and practiced in some of the city-states of ancient Greece, notably Athens, for about a century and a half and then disappeared. The basic idea of direct or participatory democracy is that all members of the political community take part in discussion, debate, and decision making on matters belonging to the public domain without the intermediation of agents. Decisions may be reached by the collectivity by voting or by reaching consensus through deliberation. As regards who qualify as members of the political community, throughout history, different criteria for exclusion–inclusion have been employed, including those based on age, gender, race, being a free man, property ownership, and paying taxes. In contemporary times, only a moderate age requirement and some other reasonable limitations such as not being insane constitute the only restrictions to being a citizen with full rights to participate in the political debate and decision making.

Athenian Democracy. The city-state of Athens is usually cited as the place where democracy was born. Yet Athenian democracy deviated in significant ways from what we know as democracy today. Only free men of 25 years or older qualified as members of the political community to participate in town meetings or Ekklesiae where discussion about the public affairs of the community affairs took place. Decisions were made by direct vote. Forty Ekklesiae were hold during a year. An administrative council whose members were identified by drawing lots was given the responsibility of implementing the decisions arrived at these meetings.

As a normative model of self rule, Athenian democracy appears attractive. From a more contemporary empirical perspective, it may be criticized for not being sufficiently inclusive since it allowed only free men of 25 years or older to participate in decision making about the political affairs of the community, leaving slaves and women out. An empirical evaluation, to the extent this is possible in view of scant information, reveals that only a small minority of citizens actually attended these meetings. In fact, meeting places were not designed to accommodate large crowds of citizens. Leaders with oratorical and organizational skills were often capable of determining both the agenda and the decisions that were eventually made. Inevitably, the Athenian system was inefficient because it took much time to produce a decision. And finally, since attendance was often limited and the composition of those who attended the meetings varied, policy coherence and continuity were difficult to achieve. Perhaps it should not be surprising that neither Plato nor Aristotle or Thucydides express admiration for the Athenian system whose legend has come to be so venerated among the later exponents of democracy.

What have we inherited from Athenian democracy? Two ideas may have been important: first, the idea of a citizen that has approximately an equal standing and voice against the state; and second, the rotation of civic responsibilities among citizens (Tilly, 2007, pp. 26–27). Many problems encountered in defining and implementing democracy in contemporary times were also anticipated in the Greek experience.

Contemporary Varieties of Direct Democracy. The Athenian democratic experiment disappeared after failing to meet external challenges. The notion of direct democracy has been practiced, however, in later times in two different ways. First, in local government, it has been adopted by small communities in the form of town meetings. Second, in some representative democracies, the mass electorates have been asked to make policy decisions by means of referenda.

Although the practice has been declining, direct democracy, sometimes also called *primary democracy*, where all members of the political community come together to decide on public matters pertaining to their town, continues to be practiced in parts of New England (town meetings) and in Switzerland (*Landgemeinde*). Direct, participatory democracy as a comprehensive system of rule appears not to be suited for communities with populations more than a few thousands or for complex tasks that contemporary governments need to address.

Another direct democratic practice with an increasing frequency in recent years is the holding of public referenda, asking the voters to determine policy choices. The more widely used form of *referendum democracy* is for the government to submit a policy question to a public referendum either on its own or because the laws require it. The less widely practiced form, known as the *citizen initiative*, allows citizens to collect a required number of signatures in order to place a proposal on the ballot. The initiative is used at all levels of government in Switzerland, in a substantial number of American states, as well as in Italy (Ian Budge, 1996, p. 85).

Direct Democracy in the Future. Technological developments appear to have opened the way for new possibilities for practicing direct democracy. With Internet service reaching a growing number of homes, it may be possible to consult voters about their preferences regarding major policy questions. The practice of such tele-democracy may lead to more widely approved policies. It has also been suggested it may even be possible in the future to realize interactive communications and create a modern version of Athenian democracy, so to speak, a cyber-democracy.

Liberal Representative Democracy

After the disappearance of Athenian democracy, there was an interim of nearly 2,000 years before the idea reappeared but this time in liberal representative form. The new system involved the placement of individuals into public decision-making roles by popular elections, thereby conferring legitimacy on the authority of those who were elected. Liberal democracy, as the name already suggests, is a synthesis of two underlying concepts: liberalism and democracy. Liberalism, although attributed changing meanings in different contexts and time points, was a reaction to the power of absolute monarchs and an interventionist church in all aspects of community and personal life. Arguing for tolerance, reason, and freedom of choice, liberalism aimed to create a uniquely private sphere for the individual in which neither the state nor the church could intervene. Placing the individual in the center of their thinking, the Liberals advocated limiting the powers of the state through constitutions and defended private property and the market economy as a means by which the individual's interests would be served (David Held, 1996). In liberal representative democratic thought, the market constituted an arena where individuals pursue their personal gain while politics served as the domain in which the interests or the common good of the community would prevail.

The idea of representative government appears to be a natural outcome of the concern of liberal philosophers with the protection of the individual against the state. This would be achieved by giving the citizens the power to choose those who govern them through periodic elections and by holding those elected accountable for their decisions and actions within the context of a constitutional system. The election of representatives to do what the citizens were expected to do on their own in a direct democracy appeared to provide the solution to the question of how large number of citizens could still rule themselves. The questions of who comprised the citizens and the conditions necessary for them to exercise the right to choose those who shall govern them continue to constitute critical questions in societies that are in the process of democratizing and, on occasion, in societies that are already democracies today. While universal suffrage appears to have acquired reasonably universal acceptance, the conditions under which political competition is to take place, such as freedom of association, information, and expression, present both existing democracies and societies that aspire to develop into liberal representative democracies with dilemmas of how to balance authority and the coercive capabilities of the state and the collective interests of the community with the liberties of the individual.

A multiplicity of social, economic, and political processes led to the emergence of liberal representative democracy as ideology. By the 19th century, aided by growing commerce and industrialization, European kings had succeeded in unifying small political entities into nation-states, creating political communities much larger than those that had existed before. The industrial revolution had also produced social classes, most notably the bourgeoisie and the workers that rejected the absolute power of the monarchs and searched for ways to limit the scope of their political decisions and influence their content. Factors such as urbanization, factory production, and new forms of communication and transportation provided opportunities for self-organization not possible in earlier times. The philosophy of liberal representative democracy developed, in this historical context, as the framework through which power could be shared between monarchs, aristocrats, and the mass publics and could gradually be transferred to the latter. The size of nation-states, in terms of population and the geographical space, necessitated an arrangement whereby public decisions would be made by relatively few people that represented the many who had elected them.

The adoption and consolidation of liberal representative democratic philosophy and arranging political systems along democratic lines did not become a fully established reality even in Western

Europe until after the World War II. Apart from Spain and Portugal, which were ruled by the remnants of fascist dictatorships of the interwar period, and Greece and Turkey, which were plagued by occasional military rule, other members of the Atlantic Community, calling themselves the "Free World," possessed regimes based on liberal democracy. The adoption of liberal representative democratic systems in the world received an initial boost from the replacement of the authoritarian Spanish, Portuguese, and later some Latin American systems by democracies. But the most significant change came with the demise of the Warsaw Pact first and then the Soviet Union. The Baltic States and many of the East European countries managed quickly to convert themselves into regimes based on liberal representative democracy. This opened the way for democratization in other parts of the world and created pressures on democracies to improve their performance.

Empirical Typologies and Types

We may discuss in normative terms what attributes a liberal representative democracy ought to have, what the proper basis of authority is, and what purposes it may serve. But in determining whether or not a particular system is a liberal representative democracy and if so, what type of a democracy it is, an empirical and descriptive approach is needed. Many scholars have, in fact, tried to develop an empirically testable set of criteria—that is, a procedural or operational definition—to distinguish political democracies from nondemocracies and then offer an empirically derived typology of democracies. A survey of democracy literature shows that a significant number of procedural elements have been proposed that identify a system as a democracy. These have included the following:

1. the right to vote, equality in voting, women's right to vote, elimination of property and wealth qualifications for voting, and inclusion of all adults;
2. the right to be elected;
3. the right of political leaders and political parties to compete for support and votes—a competitive party system;
4. periodic free and fair elections, secret ballot, and absence of massive fraud;
5. freedom of association;
6. freedom of expression;
7. freedom of information, the existence of alternative sources of information, and opportunities to learn about different policies;
8. institutions for making public policies depending on votes and other expressions of preference, responsibility of all power holders to the electorate, and opportunities for effective participation;
9. constitutions explicitly describing and limiting the authority of the power holder, institutional checks to prevent elected leaders from governing arbitrarily and without restraint, and the presence of accessible procedures for protecting the liberties of citizens;
10. control of the agenda by the elected officials;
11. the ability of elected officials to exercise their constitutional powers without being significantly constrained by unelected officials, such as bureaucrats and members of the military; and
12. a self-governing polity possessing the ability to act independently of the constraints of an overarching political system, and minimum consensus or support among the general public for values such as respect for the rights of others and tolerance.

As is evident, other elements may be added to this list, some may be combined under more general headings, and some may be broken down further. Some have also argued that if only some of these procedural elements are present, a system could be classified as a lesser or diminutive type of democracy, while others have insisted that a system is either a democracy or a nondemocracy. Diminutive types are covered in hybrid regimes in another entry in this encyclopedia.

Classifying Democracies

Among systems that are classified as political democracies, there are significant differences in terms of how political institutions are organized and how they operate. Since many of the world's democracies were initially located in Western Europe and North America or in countries that had been settled by West Europeans, initial attempts at classification focused on the differences between how democratic systems operated in English-speaking countries and countries in Continental Europe. Such geographically based designations were, however, gradually changed since democracies not only operated in other geographies but also seemed to be spreading to different parts of the world.

Majoritarian Versus Consensus Models

The departure from geography-based classifications led to the identification of two basic types: *majoritarian* and *consensus* models of democracy. To the related questions of who shall govern and whose interests the government will respond to when voter preferences diverge, two answers were offered: (1) the majority of the people or (2) as many people as possible rather than a simple majority (Arend Lijphart, 1999). If a majority, however small, was considered sufficient for governing and making rules and policies, then such a democracy was a majoritarian democracy. Since the British *parliamentary democracy* constituted the foremost example of this type, the model was frequently referred to as the *Westminster model of democracy*, though this was but one example of a majoritarian system.

In some societies, the achievement of a majority was seen as a minimum condition for governing while building as big a majority as possible was accepted as a political goal. A democracy in which an attempt was made to build the largest possible majority was referred to as the consensus model. Institutional arrangements, political practices, and patterns of interaction among political actors differed in the two models. For example, whereas executive power was concentrated in one party and the cabinet in the Westminster model, leading to the domination of the system by the executive, power was shared in broad coalition cabinets producing a government–legislature balance in the consensus model. Majoritarian democracies were usually characterized by a two-party government, consensus models by a multiparty government. The former generally used the first-past-the-post voting system, as opposed to the prevalence of proportional representation in the latter.

Efforts to develop this initial classification eventually led to the development of a more refined typology deriving from the cross-tabulation of two axes: the structure of society and elite behavior. The structure of a society could be homogeneous or plural, while the behavior of the elite could be coalescent or adversarial (Lijphart, 1977). The combination of the two axes produced four types of democracy:

1. plural-coalescent: consociational democracy,
2. plural-adversarial: centrifugal democracy,
3. homogeneous-coalescent: depoliticized democracy, and
4. homogeneous-adversarial: centripetal democracy.

Consociational and Centrifugal Democracies

In the case of societies with socially heterogeneous populations, whether the political elite pursues coalescent or adversarial behavior appears to make a determining difference. Coalescent behavior produces a *consociational democracy*, whereas adversarial behavior leads to a *centrifugal democracy*. In consociational democracies, political leaders strive to achieve the largest majority possible, incorporating the support of as many groups and citizens as they can persuade. The preference of the political elite to enlist the participation of larger numbers than needed is a way of controlling the centrifugal tendencies that are usually present, even if not always inherent, in plural societies. To achieve a substantial majority at the societal level, leaders of different segments of the population work together to mobilize majorities in their respective communities. They also respect mutual vetoes. In this way, political power is dispersed and shared. To cement the system further, public funds are allocated to different groups comprising the plural society on a proportional basis and each group is granted extensive autonomy to organize its own affairs.

If the political elite in a society with a heterogeneous population pursues adversarial behavior, if they insist on considering the achievement of a bare majority sufficient to produce decisions that are binding for all (i.e., if they pursue a policy of strong majoritarianism), those finding themselves frequently or permanently in the minority may find the outcome unacceptable. In such a contingency, the tendency would be for each group to form a coalition in which it would always aim to be in the winning camp, directing it to undo any coalition where it is not included or alternatively opt out of a system that consistently fails to serve its interests, hence the centrifugal democracy. For example, one way to cope with inability to form stable and working coalitions in the face of centrifugal tendencies particularly (but not only) in a presidential system is for an elected leader to opt for *delegative democracy*. Generally, consociational democracies are rather stable, while centrifugal democracies are exposed to intensive tensions, instability, and the possibility of malfunctioning and breakdowns.

Depoliticized and Centripetal Democracies

To the extent societies and their politics are always changing, systems of specific countries may move from one type to another. For example, during World War II, many democracies in Western Europe turned into *depoliticized democracies*. All political parties were united around the fundamental goal of winning the war, while politics was perceived less in government versus opposition terms and more in terms of maintaining the unity of all political actors to achieve the common objective of victory. There may also be occasions when temporary conditions encourage the forming of a coalition between government and the opposition, such as the one between the Christian Democratic Union and the German Social Democratic Party under Kurt Georg Kiesinger in Germany (1966–1969), which led to the practice of depoliticized politics for a limited period of time.

Depoliticized democracy and centripetal democracy are distinguished by the fact that in the latter, politics is conceptualized much more as a competitive game between government and the opposition, whereas depoliticized democracy is characterized by the minimization of conflict and a sustained effort toward achieving a broadly based consensus. The British system is given as the example of a centripetal democracy. In the U.S. system, the majority party in the Senate and in the House will generally try to ensure that its members vote with the party on particular pieces of legislation. In centripetal democracies, the competition to achieve a minimum majority (strong majoritarianism) does not threaten the unity of the system or the integration of rival parties into it. The alternation of majorities between elections reinforces the desire of competing parties to win the elections, inviting them to adopt positions that will appeal to large groups of voters and drawing them toward the center—hence the designation as centripetal democracy. The homogeneity of social structure stands in the way of political fragmentation and the consequent emergence of permanent minorities that may sometimes even desire to break off from the system. Strongly adversarial conceptualization of oppositions is further eroded by two other factors. First, in many economically advanced societies, ideological, religious, and to a more limited extent racial tensions have receded to the background in recent decades, leading voters to fail to perceive major differences between competing parties. Therefore, a change in the governing party is not seen to be particularly critical. Second, all major interests in society find opportunities to be represented in decision making through a variety of mechanisms. These points will be further elaborated in the models of *polyarchy* and *functional democracies*.

Polyarchy

The critical problem that needs to be overcome in sustaining a democratic system and ensuring its stability appears to be that no group be cornered into a position of a permanent political minority—that is, be placed permanently in an inferior status under the mercy of a permanent majority. Such an outcome may be achieved either by shifting majorities or by ensuring that the majority is as large as possible. The latter option is associated with the consociational democracy that is usually found in plural societies. There are many societies that are neither plural nor consciously subscribe to the consensus model. What is the mechanism in such societies for ensuring that the same majority does not prevail over the same minority on a permanent basis? The simple answer is that voters change their preferences. In a democratic society where people can form associations, express their thoughts, and reach alternative sources of information freely and where free and fair elections are held regularly, voters have the opportunity to replace the majority in power by another majority. But, this explanation does not rule out the possibility that some groups would be left permanently out of the decision-making process.

The way out of this predicament, Robert Dahl argued (1971), is the recognition of how contemporary democracies operate in fact. Although political competition appears to be dominated by political parties, there is also a proliferation of organized interest groups. The political competition in a democratic society does not simply produce a majority and a minority. Rather, a plurality of actors compete and cooperate with each other, usually forming issue-based majorities, to produce policy decisions. Different coalitions form behind different policy choices in an environment in which power is dispersed among various actors in society. What obtains is not majority rule as such but a rule of different coalitions of minorities. These minorities negotiate, bargain, compromise, and form alliances and coalitions to produce outcomes that they desire. Dahl called his explanation of how democracy works *polyarchy*.

Its proponents argue that polyarchy is an empirically derived description of how democracies function. It has been criticized on the ground that it relies to a large extent on the study of American experience. It assumes that organized interests prevail in a society, that these represent a significant segment of society, and that through competition they are capable of controlling each other such that none is able to prevail permanently over others. These assumptions require empirical scrutiny particularly outside of the United States and specifically in European and non-Western environments. Some sample questions would include whether all groups in society are equally capable of organizing themselves into effective organizations, whether they have equal resources to influence the political process, and whether some groups get much of what they want much of the time while others get hardly anything at all.

Polyarchy involves a permanent process of building coalitions among different groups to form majorities. Each majority comprises a different configuration of groups. This would lead actors to confine themselves to the achievement of bare majorities rather than pursuing the consensus model and trying to build as large a majority as possible.

Functional Democracy

The polyarchy model assumes that multifarious interests will organize and compete with each other in affecting public policies in their favor, with none getting all of what it wants. There does exist a rather different way of linking interest groups to the democratic political process: that of corporatism. Although *corporatist* or *functional democracy* may take many forms, its basic idea is to incorporate major organized interests (usually each represented by a single organizational entity) into the policy-making and implementation processes. Corporatist arrangements, it is sometimes argued, are more effective in ensuring that the interests of all major groups are represented in politics. Yet a number of questions remain. First, how can we be sure that all major interest groups are included in this arrangement? This question becomes all the more important since any corporatist system would tend to work in favor of those groups that have existed for a very long time and at the expense of those that are newly emerging. Second, we may ask, are the interest groups that are linked with the government run democratically themselves? Again, it is important to know the answer not only because there may exist different interpretations of what is in the interest of a particular group or what its interests are, but also because group leaders often develop an interest in sustaining themselves in office at the expense of the groups that they are supposed to represent and whose interests they are supposed to protect. Third, what concessions do the leaders of organized interests have to make to maintain their favored status vis-à-vis the government? Implicit in the question is the possibility that the leaders of interest groups may simply ease into becoming agents of government policy, transmitting, explaining, defending, and even helping implement it. Fourth, what effect do the corporatist arrangements have on other democratic institutions? For example, do these arrangements undermine respect for institutions such as political parties that have traditionally served as a channel for managing conflicting interests? These questions point to some of the potential shortcomings of functional democracy.

Delegative Democracy

In examining how democracies functioned in Latin America, Guillermo O'Donnell observed that while these systems met the criteria for being designated "democratic," they failed to display many of the characteristics of liberal representative democracies because they operated on the premise that the winner of the national election acquired the right to govern as he or she sees fit. The systems under study are presidential systems, where apparently the president feels that for a constitutionally defined period of time between elections, he or she is at liberty to rule mainly according to his or her preferences, only constrained by the distribution of power in society. O'Donnell termed these systems *delegative democracies*.

In systems where delegative democracy prevails, accountability runs vertically, whereas in institutionalized representative democracies, there is, in addition, horizontal accountability. Vertical accountability denotes electoral accountability. Horizontal accountability, on the other hand, refers to the presence and effectiveness of a network of autonomous institutions that exercise oversight on government officials to ensure that they discharge their responsibilities properly or be subject to sanctions. Presidents in delegative democracies treat institutions of horizontal responsibility as unwanted impediments to the performance of their duties as the custodian of national interest. They accept, however, the existence of rival parties and a reasonably free press as normal.

Delegative democracy is a less liberal form of democracy than a fully representative liberal democracy. It is strongly majoritarian with elections that give powers to a president who claims to serve as the "interpreter of the high interests of the nation." Although they are not institutionalized, delegative democracies may be enduring.

Parliamentary and Presidential Democracies

When one focuses on the institutional arrangements within the liberal representative democracies, the most common distinction is between parliamentary democracies and presidential ones. The *presidential arrangement* was the product of the American Revolution. Not having a fully functioning parliamentary system before them that they could choose to emulate, the leaders of the American Revolution converted the institution of the Absolute Monarch to a president who was elected for a defined period of time and with limited powers. *Parliamentary democracy*, on the other hand, evolved in Britain and Continental Europe as a result of the rise of new social classes during the industrial revolution that claimed a share in ruling society. The conversion of consultative mechanisms among the king, the nobility, and sometimes the church into parliaments whose members were elected through a competitive process was achieved often through a combination of violent change and peaceful transfer of power.

Presidential democracy as initially developed in the United States was part of a constitutional system based on the separation of powers as well as building checks and balances between the three branches of government. The president was elected by an electoral college whose members were elected explicitly to elect a president. The members of the legislature, the U.S. Congress, on the other hand, were elected directly through general elections. The president, elected independently, did not need to maintain the confidence of the legislature to stay in power and served a 4-year term. Initially by convention, much later by constitutional requirement, his tenure was limited to two terms. The American presidential system became more democratic over time. The adoption of a Bill of Rights, the abolishment of slavery after a destructive civil war, and the extension of franchise to Blacks and later to women constituted some of the landmark developments in this evolution.

The presidential system was adopted by almost all Latin American countries, but in none did it function as democratically as in the United States. In many instances, the presidential system constituted a method of promoting full or semidictatorships. Many Latin American presidential democracies today (not all are necessarily democratic) continue to exhibit the characteristics of delegative democracies.

Parliamentary democracy, a must for royal systems in which the office of the chief executive and the head of state is hereditary, is mainly a product of the 20th century. While some parliamentary systems were in operation intermittently in some European countries and somewhat more regularly in Britain, they were neither fully democratic nor consolidated. Most European parliamentary democracies did not extend suffrage to women until after World War I. Introduction of universal suffrage destabilized many of them, leading some, such as the Weimar Germany, Spain, Italy, and Austrian systems, to collapse while producing deep political polarization in such countries as Belgium, France, and Finland.

Parliamentary democracy is characterized by a fusion of powers in that the government is generally formed among parties in the parliament. The prime minister is almost always required to be an elected member of parliament, while other ministers are also usually deputies even if that is not always a legal requirement. The government is required to enjoy the confidence of a parliamentary majority to stay in office. The parliament, on the other hand, scrutinizes the government by questions, inquiries, committee hearings, interpellations, and motions of censure or no confidence. But, not unlike the presidential variant, parliamentary democracies exhibit significant variety among themselves.

Presidential and parliamentary democracies differ from each other in important ways. As already said in part, legislatures may not dismiss a president (except in cases of corruption, mental incompetence, etc.), while parliaments, although subject to some limitations, may dismiss a government by a vote of no confidence. In parliamentary democracies, governments may choose the timing of the elections, while election dates are more likely to be fixed in presidential democracies. Almost universally, presidential democracies have term limits for the president because he or she wields so much power. One is more likely to find multiparty systems in parliamentary than in presidential democracies. Although it has been suggested that presidential democracies enjoy strong, personalized leadership, the rise of party leaders to unequal supremacy in parliamentary democracies has also produced a similar type of leadership. Furthermore, many parliamentary democracies have evolved in the direction of prime ministerial government where the former many appoint and dismiss ministers pretty much as if he or she were a president.

There is a hybrid institutional arrangement that also deserves brief mention: the *semipresidential democratic system*, an arrangement that is most closely associated with France. In these systems, the elections of the president and the members of parliament usually take place in coterminous but distinct elections. The timing of the elections is fixed. In the French case, to ensure that the mandate of the president is backed by a majority, a runoff election is required among leading contenders. There is a division of responsibilities between the president and the prime minister, not very precise, regarding domains in which they are to be active. As in the case of presidential democracies, in semipresidential systems, the majority in the legislature and the partisan affiliation of the president may be at variance, necessitating cohabitation.

Critical Approaches: Democracy in Theory and in Practice

The diffusion of liberal democratic ideologies and the prevalence of democratic systems in the world have produced two contradictory outcomes. On the one hand, the crumbling of nondemocratic systems has given satisfaction to those that have been ruled by democratic regimes as well as a sense of confidence that democracies have outlasted what appeared to be powerful authoritarian regimes. On the other hand, democracies have come under greater scrutiny as scholars have begun to take a closer look at the imperfections of democracy as an ideology and as an operating political system.

Criticism of Democracy in Operation

Some problematical outcomes of democratic systems in operation have been observed and criticized by different observers. One relatively frequently made observations is that the liberal representative model has evolved into *party democracy* in which representatives are obliged to serve the parties under whose auspices they have been elected; therefore, they are neither able to express their own judgment nor represent without constraints the views of their own constituency. Parties, on the other hand, publicize their policy priorities in the electoral campaign and then proceed to implement them once in office. Parties, however, are in fact generally directed by their leaders. In this way, democracy in operation has become *plebiscitary leadership democracy*, an exercise in which the voters have periodic chances to express or deny confidence in the political leadership that runs the government (Held, 1996). Furthermore, the parties in power may represent only a minority of the voters and may use their being in office as a resource for extending their rule.

Approaching democracy in operation from another angle, it has been suggested that contemporary democracy is *elitist democracy* or that it is characterized by *democratic elitism* because key political decisions are made by small minorities. This is seen as a necessity deriving from the complexities of decisions, the time frame within which decisions have to be made, and the voluminous topics on which decisions have to be rendered. In this process, the role of the voter is no more than choosing among rival teams of elites. Even more problematical, the voters are often not well-informed and not sufficiently committed to the fundamental values of democracy such that too much involvement on their part may become a source of instability, a threat to the functioning of the democratic system (Peter Bachrach, 1969).

Radical Democracy

Radical democracy is an imprecise term that offers criticism of the contemporary applications of democracy from a variety of perspectives. Many persons whose names are associated with radical democracy are not interested in the replacement of the liberal and pluralistic democracy with another type, rather, they argue that it may be deepened and expanded to cover many areas of public life. One example is a deliberate effort to incorporate into the public all potential voters by enfranchising them so as to curb the ability of majorities to use their position to maintain their dominance of society. Another example is the struggle to undo the relations of subordination deriving from differences in gender, race, and sexual preference in addition to those deriving from the political economy. In all these instances, it is argued that the expansion of liberty and equality to all aspects of public life within the framework of political liberalism will enhance democracy and protect the rights of the individual against the tyranny of the majority. To conclude, the various proponents of radical democracy argue that contemporary applications of liberal democracy fail in being sufficiently inclusive to bring in all segments of the population into the political process. They further argue that current liberal democracies do not sufficiently extend liberal democratic applications to all areas of public life.

Deliberative Democracy

The prevalence of liberal representative democracy as the normative basis for the functioning democracies in the world today has led some analysts to focus on some of the problematical aspects and outcomes and proposals to remedy them. The best known school of thinking that has come under different names such as *communicative* or *discursive*, but most frequently as the *deliberative theory of democracy*, offers a normative framework for addressing some of the problematical aspects of liberal democratic theory.

Proponents of deliberative democracy point out that the two ideas of liberalism and democracy that are brought together in liberal democracy in fact harbor conflicts. While liberalism dwells on pluralism, individualism, and freedom, critical democratic principles include unity, community, and equality. First, there is a conflict between the liberal emphasis on pluralism and the basic need for the basic social and political unity of democratic society. Second, while individualism is ontological to liberalism, democracy assumes communitarianism. Finally, freedom, so fundamental to liberal thinking, and the principle of equality, so fundamental to democratic thought, may be seen as being pitted against each other.

These conflicts, in fact, are not conciliated in liberal democracy because while there may be argument and bargaining, the decision is formed by voting, which is an aggregative act, not one of reasoning in which all members of the community are persuaded by argumentation and discussion. Therefore, a decision that emerges as a result of voting does not reflect a consensus among all citizens. Deliberative democracy argues that free and open debate among equal citizens should continue until the emergence of a community consensus. In this process of what may be called practical reasoning, participants offer ideas and proposals for how to best solve problems and meet the needs and so on of the community. Each presents arguments to persuade the others to accept his or her proposals, eventually leading to a consensus. In this model, democratic process is primarily a discussion of problems, disagreements, expressed needs, and interests where, through dialogue, each tests and challenges the proposals and arguments of others.

For a deliberative democracy to operate, citizens must share or demonstrate a commitment to solve their collective problems through public reasoning and accept the framework for public deliberation as legitimate. Jürgen Habermas has argued that the successful operation of the deliberative democratic procedure is based on the following postulates:

1. Deliberation is in argumentative form. There is a regulated exchange of information among parties who introduce and critically test proposals.
2. All those who are affected by the decision have an opportunity to take part in the debate. In principle, no one may be excluded from the deliberations that are public.
3. The deliberations are free of external and internal coercion that undermine the equality of participants.
4. The deliberations aim to achieve rationally based agreement. They may be discontinued and resumed any time.
5. It is expected that the predeliberation attitudes and preferences of parties will be changed through arguments.

It is evident that deliberative democracy does not offer a practical alternative to the operation of contemporary democracies. Its application in societies with large populations where the governments are expected to address many problems simultaneously is not a realistic possibility since the time needed to implement the process would be unacceptably long and near-full consensus might be impossible to attain in many situations. Nevertheless, the proponents of deliberative democracy point to some problematical aspects—that is, contradictions and weaknesses in contemporary applications of liberal democracy. Reaching decisions by majority vote may exclude many of those most closely affected by a decision; majorities and minorities may display a reasonable permanent character, frustrating those that are pushed to the position of a minority.

Cosmopolitan Democracy

In contrast to the radical and deliberative criticisms of contemporary liberal democracy that emphasize current shortcomings, the proponents of cosmopolitan democracy point to the direction they think democracies should take in view of the changes in the world. Noting that democratic systems get transformed and become redefined as societies evolve, as new needs appear, and as some functions and institutions no longer respond to needs and become redundant, Daniel Archibugi and David Held (1995) have pointed out that democratic systems operate under the highly questionable assumption that a government exercises sovereignty over a national political community that inhabits a delineated territory. Yet national, regional, and global connectedness belies the validity of this assumption. Therefore, there is a need to adjust democratic systems and applications to global change. The solution may be the development of a multipronged global model of democracy where innovations such as regional parliaments would produce decisions that would be recognized as sources of law; the holding of referenda across nation-states on issues such as energy policy, transportation systems, and so on; and the entrenchment of a body of individual rights with a political, economic, and social content. It may be noted, in this regard, that some basis has already evolved for the implementation of these ideas since there are already various products of regional integration in the world as well as a growing acceptance of universal standards for the observation of human rights. These may well mark the beginning of a process that will take us to a democracy that transcends the borders of the nation-states, keeping in mind, however, that there is no single inevitable way societies respond to pressures for change.

Social Democracy Versus Liberal Democracy

The discussion of the types of democracy has focused so far on institutional arrangements. Since political democracy is usually thought of as a political regime, this is natural. The content of social and economic policies that democracies have pursued, however, has also constituted grounds for distinguishing between two policy-based types: liberal and social democracies. As a political ideology, liberal democracy and social democracy are based on the two rival premises of equality of opportunity and equality of condition, respectively. Liberal democracy argues that the market constitutes the most reliable instrument for offering equality of opportunity to all citizens, leading it to advocate policies that rely on the market as the mechanism that brings solutions to social and economic problems. Consequently, it favors policies that promote the undisturbed operation of the market, such as not introducing wage and price controls, maintaining low taxes, and pursuing free trade. It expects the individual to pay fully for his or her needs, such as housing and health care. Social democracy, on the other hand, assuming a society in which members consider themselves as being socially equal, proposes the adoption and implementation of policies that aim to enhance the equality of conditions, such as a large public presence in the economy, minimum wages, unemployment insurance, social security, universal health service, and free public education, as well as other similar functions. Many political parties that pursue social-democratic goals and policies also bear the name "Social Democratic" (in capital letters).

It has been observed that originally liberal democracy evolved in societies where the middle classes were united and coherent, while the urban working class was weakly organized and divided. In contrast, social democracy emerged in areas where the middle classes were fragmented and lacked coherence, while urban working classes were capable of organizing and forcing a social compact. As some of the policy goals initially sought by social democracy (e.g., social security, minimum wages, unemployment insurance) became standard policy in contemporary democratic societies, the policy content of both liberal and social democracy has become redefined, and their differences have become less clear. Nevertheless, these two types continue to constitute one of the fundamental bases of political choice on which electoral competition is conducted in most societies ruled by democracy.

Ilter Turan
Istanbul Bilgi University
Istanbul, Turkey

See also Deliberative Policy Making; Democracy, Consociational; Democracy, Direct; Parliaments; Presidentialism; Representation

Further Readings

Archibugi, D., & Held, D. (Eds.). (1995). *Cosmopolitan democracy*. Cambridge, UK: Polity Press.

Bachrach, P. (1969). *The theory of democratic elitism*. London: University of London Press.

Budge, I. (1996). *The new challenge of direct democracy*. Cambridge, UK: Polity Press.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Dahl, R. A. (1998). *On democracy*. New Haven, CT: Yale University Press.

Elster, J. (Ed.). (1998). *Deliberative democracy*. Cambridge, UK: Cambridge University Press.

Habermas, J. (1996). *Between facts and notes*. Cambridge, UK: Polity Press.

Held, D. (1996). *Models of democracy*. Oxford, UK: Polity Press.

Hirst, P. (1990). *Representative democracy and its limits*. Cambridge, UK: Polity Press.

Lijphart, A. (1977). *Democracy in plural societies*. New Haven, CT: Yale University Press.

Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty six countries*. New Haven, CT: Yale University Press.

Mouffe, C. (1996). Radical democracy or liberal democracy. In D. Trend (Ed.), *Radical democracy: Identity, citizenship, and the state* (pp. 19–41). New York: Routledge.

O'Donnell, G. (1996). Delegative democracy. In L. Diamond & M. F. Plattner (Eds.), *The global resurgence of democracy* (pp. 94–108). Baltimore: Johns Hopkins University Press.

Sartori, G. (1987). *The theory of democracy revisited*. Chatham, NJ: Chatham House.

Saward, M. (1998). *The terms of democracy*. Cambridge, UK: Polity Press.

Tilly, C. (2007). *Democracy*. Cambridge, UK: Cambridge University Press.

von Beyme, K. (2000). *Parliamentary democracy*. New York: St. Martin's Press.

Index

Entry titles and their page numbers are in **bold.**

International Encyclopedia of
POLITICAL SCIENCE

VOLUME 3

Editorial Board

International Encyclopedia of
POLITICAL SCIENCE

VOLUME 3

EDITORS

BERTRAND BADIE
Sciences Po, Paris, France

DIRK BERG-SCHLOSSER
Philipps-Universität Marburg, Germany

LEONARDO MORLINO
LUISS Guido Carli, Italy

SAGE reference

Los Angeles | London | New Delhi
Singapore | Washington DC

IPSA AISP

Developed in partnership with the
International Political Science Association

Los Angeles | London | New Delhi
Singapore | Washington DC

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
33 Pekin Street #02-01
Far East Square
Singapore 048763

Publisher: Rolf A. Janke
Assistant to the Publisher: Michele Thompson
Developmental Editor: Diana E. Axelsen
Production Editor: Tracy Buyan
Reference Systems Manager: Leticia Gutierrez
Reference Systems Coordinator: Laura Notton
Copy Editor: QuADS Prepress (P) Ltd.
Typesetter: C&M Digitals (P) Ltd.
Proofreaders: Kristin Bergstad, Kate Peterson, Sally Scott
Indexer: Mary Mortensen
Cover Designer: Janet Kiesel
Marketing Manager: Kristi Ward

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

International encyclopedia of political science / edited by Bertrand Badie, Dirk Berg-Schlosser, Leonardo Morlino.

8 v. cm.
A SAGE Reference publication.
Includes bibliographical references and index.

ISBN 978-1-4129-5963-6 (cloth)

1. Political science—Encyclopedias. I. Badie, Bertrand. I
I. Berg-Schlosser, Dirk. III. Morlino, Leonardo

JA61.I58 2011
320.03—dc22 2011009801

11 12 13 14 15 10 9 8 7 6 5 4 3 2 1

Contents

3

List of Entries

Reader's Guide

Comparative Politics, Theory, and Methods

Case and Area Studies

Case Studies
Caudillismo
Communist Systems
Comparative Methods
Comparative Politics
Cross-National Surveys
Democracy: Chinese Perspectives
Democracy: Middle East Perspectives
Democracy: Russian Perspectives
Fascist Movements
Multiculturalism
Populist Movements
Postcommunist Regimes
Regional Integration (Supranational)

Subnational Governments

Autonomy, Subnational
Decentralization
Federalism
Governance
Governance, Urban
Local Government
Peace
Regionalization
Self-Determination

Democracy and Democratization

Accountability
Accountability, Electoral
Accountability, Interinstitutional
Change, Institutional
Citizenship
Civil Service
Coalitions
Collaboration
Colonialism
Competition, Political
Conditionality
Constitutional Engineering
Constitutionalism
Corruption, Administrative
Credible Commitment
Democracy, Direct
Democracy, Quality
Democracy, Types of
Democracy: Chinese Perspectives
Democracy: Middle East Perspectives
Democracy: Russian Perspectives
Democratization
Developing World and International Relations
Development, Political
Development Administration
Empowerment
Federalism
Foreign Aid and Development
Governance
Governance, Good
Groupthink
Human Development
Liberalization
Modernization Theory
Monarchy
Nation Building
Opposition
Peasants' Movements
Pluralist Interest Intermediation
Postcolonialism
Postmaterialism
Representation
Responsibility
Responsiveness
Responsiveness of Bureaucracy
Rule of Law
Self-Determination
Semipresidentialism
State Collapse
State Failure
State Formation
Sustainable Development
Traditional Rule
Transition
Transitional Justice

Decision Making in Democracies

Cost–Benefit Analysis
Delegation
Deliberative Policy Making
Election by Lot
Election Observation
Election Research
Elections, Primary
Elections, Volatility
Electoral Behavior
Electoral Campaigns
Electoral Geography
Electoral Systems

Epistemological Foundations

Gender and Race/Ethnicity

International Relations

Peace, War, and Conflict Resolution

Political Economy

Welfare Policies
Welfare State
World Bank
World Trade Organization (WTO)

Political Parties

Christian Democratic Parties
Cleavages, Social and Political
Communist Parties
Conservative Parties
Green Parties
Liberal Parties
One-Party Dominance
Parties
Party Finance
Party Identification
Party Linkage
Party Manifesto
Party Organization
Party System Fragmentation
Party Systems
Social Democracy
Socialist Parties

Political Philosophy/Theory

African Political Thought
Anarchism
Charisma
Communism
Communitarianism
Conservatism
Constitutionalism
Contract Theory
Democracy, Theories of
Discursive Institutionalism
Ethics
Fascism
Fundamentalism
Greek Philosophy
Idealism in International Relations
Liberalism
Liberalism in International Relations
Libertarianism
Liberty
Maoism
Marxism
Mercantilism
Nationalism
Neoliberal Institutionalism
Neoliberalism
Normative Political Theory
Normative Theory in International Relations
Pacifism
Pluralism
Political Class
Political Philosophy
Political Psychology
Political Theory
Postmodernism in International Relations
Realism in International Relations
Revisionism
Rights
Secularism
Socialism
Stalinism
Statism
Theocracy
Utilitarianism
Utopianism

Equality and Inequality

Cost–Benefit Analysis
Equality
Equality, Political
Justice
Marxism
Racism
Socialist Systems

Formal and Positive Theory

Bureaucracy, Rational Choice Models
Game Theory
Principal–Agent Theory
Prisoners' Dilemma
Rational Choice
Social Choice Theory
Veto Player

Theorists

Durkheim, Émile
Hegel, Georg Wilhelm Friedrich
Hobbes, Thomas
Kant, Immanuel
Locke, John
Machiavelli, Niccolò

Political Sociology

Public Policy and Public Administration

Public Policy

Public Administration

Qualitative Methods

Quantitative Methods

Religion

Democracy: Chinese Perspectives

As a universal idea today, democracy has different meanings in various countries. The Chinese perspectives of democracy are a product of not only Western cultural diffusion but also ideas that were shaped by the country's history and political reality. Indeed, the vision of the rule by the people came to China from abroad. However, the thought of the rule for the people (*minben*) has deep historical roots and now takes a central position in Chinese official ideology. As the Communist Party of China (CPC) is defined as the only ruling party constitutionally, multiparty competition is beyond the imagination of most Chinese due to their value orientation and political calculation. In other words, Chinese history, ideological context, and constitutional framework as well as Western cultural influences have set up parameters for the Chinese perspectives of democracy. In the following, both the official view of Chinese democracy and some contending voices concerning other and more recent positions are examined.

Socialist Democracy With Chinese Characteristics: The Official View

China officially claims to unswervingly develop socialist democracy as an indispensable goal of socialism with Chinese characteristics. According to the official doctrine, the Chinese socialist democracy must integrate "the leadership of the Communist Party, the position of the people as masters of the country (*renmin dangjia zuozhu*), and the rule by law (*yifa zhiguo*)." Such a trinity, reflecting an ideological mix of guardianship and populism, demonstrates the basic notion and principle of democracy in the Chinese official view. From the perspective of the Chinese authorities, the ideal type of democracy is the socialist democracy with Chinese characteristics. Its basic form is people's democratic dictatorship, a phrase incorporated into the Constitution of the People's Republic of China (PRC) by Mao Zedong, the founder of the republic. Such a phrase, derived from the dictatorship of the proletariat, is notable for being one of the few cases in which the term *dictatorship* is used in a nonpejorative manner. The premise of the people's democratic dictatorship is that the party and state democratically represent and act on behalf of the people but possess and may use dictatorial powers against reactionary forces.

Such a form of democracy means to ensure democratic rights for people while exercising dictatorship in relationship to people's internal enemies. Implicit in the concept of the people's democratic dictatorship is the notion that dictatorial means are a necessary evil and that, without a dictatorship, the government may collapse and create a situation that is worse than the dictatorship. In the post-Mao era, however, "dictatorship" has been mentioned more rarely, and by contrast, "socialist democracy" has gained more prominence. As the former Chinese leader Jiang Zemin argued, the essence of people's democratic dictatorship is people's democracy. The 17th CPC National Congress in 2007 proclaimed that people's democracy is the lifeblood of socialism. Etymologically, democracy means people's rule. By using the tautology "people's democracy," the CPC attempts to conceptualize socialist democracy and distinguish it from liberal democracy, which is usually regarded as Western democracy or capitalist democracy. According to the Chinese official doctrine, capitalist democracy or bourgeois democracy is only a democracy for a few people—the bourgeoisie, while the people's democracy or socialist democracy means a democracy for the overwhelming majority of the people. Theoretically, the socialist democracy with Chinese characteristics aims to develop the people's democracy to a high degree; as Party General Secretary Hu Jintao put it, the party has been consistently pursuing the goal of developing socialist democracy, and the essence and core of socialist democracy are that the people are masters of the country.

The "people's democracy" in China, therefore, carries a strong flavor of direct democracy as opposed to liberal democracy. At least, democracy in the Chinese context is, to some extent, a mixture of direct democracy and representative democracy. According to the Chinese official proclamation, the CPC must ensure that all power of the state belongs to the people, expand the orderly participation of citizens in political affairs at each level and in every field, and mobilize and organize the people as extensively as possible to manage state and social affairs as well as economic and cultural programs. In the official conception of people's democracy, democratic election, decision making, administration, and oversight are the four important elements featured by the spirit of direct democracy, whereas the ideas of democratic decision making and democratic administration both presume that people are capable of getting involved in public administration and policy making directly without relying on their representatives. This ideological inclination is consistent with the populist view of politics, which assumes that each person has an equal say on all policy issues at all times. Actually, for most Chinese, including many leaders, democracy means people's rule, government by a mass of people, "the position of the people as masters of the country," or even mob rule and anarchism.

By contrast, the idea of socialist democracy with Chinese characteristics, or people's democracy, carries a legacy of people's democratic dictatorship. Although the Chinese leadership has downplayed the concepts of class struggle and social revolution since China's reforms and opening up in 1978, it still maintains concepts such as "antagonistic forces" (*didui shili*) and "contradiction between people and its enemy" (*diwo maodun*) in the official language. Any hostile forces are subject to people's dictatorship, without the freedom of enjoying people's democracy. A truly liberal democracy is applicable to all citizens (except for criminals) in a community.

Systems

To support people's democracy, the party and state have established a series of political systems, including the system of people's congresses, the system of multiparty cooperation and political consultation under the leadership of the CPC, the system of regional ethnic autonomy, and the system of self-governance at the primary level of society.

The system of people's congresses, as China's fundamental political system, is an organizational form of the state power in China. According to China's constitution, the power in the People's Republic of China belongs to the people, and the organs for the people to exercise state power are the National People's Congress (NPC) and local people's congresses at all levels. The NPC and local people's congresses are established through elections, and they are responsible to and supervised by the people. State executive and judicial organs at the different levels are created by, responsible to, and supervised by the people's congresses. The NPC is the organ with supreme state power, and local people's congresses are local organs of state power. The formal powers of the people's congresses stipulated by the constitution, however, have been neutralized by the leading role of the CPC, which is defined as the only ruling party by the constitution and operates under the principle of democratic centralism. Despite the growing assertions of the NPC and its standing committee vis-à-vis the State Council, the Supreme People's Court, and the Supreme People's Procuratorate, the NPC's supreme powers in elections, legislation, and supervision are constrained by the Party's discipline and ideology.

Historically, the NPC was designed according to the Marxist–Leninist principle of "combining legislative and executive into one organ" (*yixing heyi*), in order to ensure power concentration and political efficiency. According to Marx, the principle of checks and balances would leave governmental branches outside the direct control of the electorate; all levels of government therefore should be a working body fully accountable to the people. Informed by Marx's idea, Lenin adopted the model of the Paris Commune and created the Soviet regime in Russia. To transform an old bourgeois parliament—"a talkative organ" in Lenin's words—into a proletarian working body, Lenin proposed that every member of the Soviet should take both legislative and executive jobs. Following the Leninist principle of democratic centralism and the Soviet model of political regime, the Chinese ruling elite denigrated the liberal proposition of separation of powers from the very beginning. For the ruling elite, institutional checks and balances among legislative, executive, and judicial branches would lead to a divided sovereignty and therefore reduce political efficiency. These are at odds not only with the ruling party's belief in democratic centralism but also with the traditional vision of the rule by virtue (*dezhi*), which prescribes a unitary and harmonious government as an ideal goal.

The multiparty cooperation and political consultation under the leadership of the CPC is a basic political system in China. Besides the CPC, there are eight minor "democratic parties" (*minzhu dangpai*) in China, which are neither the ruling parties nor opposition parties, because they do not challenge the leadership of the CPC and demand for competitive elections in China. Rather, they are called "participatory parties" (*canzheng dang*) mainly through the mechanism of political consultation, particularly via the institution of the Chinese People's Political Consultative Conference (CPPCC). The political powers and rights enjoyed by the "democratic parties" are delegated or endowed by the ruling party, rather than being based on their own electorate. Nevertheless, the people's congresses, the CPPCC, and government agencies at all levels have members from the "democratic parties" as part of their leaders. Occasionally, some members of "democratic parties" can serve as ministers in the State Council.

In addition to the system of people's congresses and the system of multiparty cooperation and political consultation, self-governance in some ethnic areas and grassroots communities normatively displays one of the features of the Chinese socialist democracy, which are called the system of regional ethnic autonomy and the system of self-governance at the primary level of society. Empirically, self-governance in Chinese villages has developed rapidly in the past years, and especially, competitive elections for village committees have been spreading all over the country.

Conceptualizing a Chinese Model of Democracy: Academic Exploration

Although the Chinese official view of democracy illustrates an outline of the pattern and framework for democracy and democratization in contemporary China, it is far from a sophisticated theory. Among others, one of the vague issues is what the "Chinese characteristics" mean within the socialist democracy advocated by the Chinese authorities. In other words, what model could China develop for democracy? With reference to such a question, Chinese intellectuals have different viewpoints under the umbrella of socialist democracy or beyond, in search of an ideal or possible democratic model for the future China.

Direct Democracy/Autonomous Democracy

The socialist democracy or people's democracy in China, as mentioned above, has an orientation of direct democracy. Accordingly, the official doctrine insists on democratic elections, democratic decision making, democratic administration, and democratic oversight, which constitute the basic framework of Chinese orthodoxy on democracy with properties of autonomy or direct democracy. Furthermore, grassroots autonomy in Chinese villages has developed rapidly since the 1990s, giving an empirical support to such orthodoxy. Resulting from this development, autonomous democracy is regarded as one major option of the Chinese democratic model.

From a perspective of the state–society dichotomy, some scholars write that the political system can be divided into state power and grassroots communities. Accordingly, the state and grassroots levels of democracy can also be distinguished; at the state level, indirect or representative democracy is employed, and at the grassroots level, direct or autonomous democracy is practiced. In China, grassroots democracy includes rural autonomy in villages, urban autonomy in communities, democratic administration in the enterprises, and urban and rural civic participation in public affairs. What is more, some scholars believe that villagers' autonomy would be a breakthrough or growing point for China's democratization. To enhance autonomous democracy, some scholars suggest expanding the scope of grassroots democracy from villages and urban neighborhood communities up to the township or higher levels.

Besides the grassroots democratic autonomy in the rural and urban communities, some other kinds of direct democracy have also emerged in recent years. Some leaders and scholars emphasize the importance of citizens' direct participation in public affairs and policy making through public hearings, participatory budgeting (open discussion on local public budgets), Internet discussions, and so on. In particular, cyber democracy has developed quite rapidly and is more prevalent compared with other forms of democracy in China, even exceeding its development in Western democracies to some extent. This is mainly because in China there are hardly any other effective channels for citizens to articulate their views except for the Internet. More and more Chinese people prefer to use the Internet as a means to express their opinions on public affairs, to influence public policy, and to oversee public officials. With the greatest population of netizens in the world, China's cyber democracy is developing ahead of representative democracy and is welcomed by most Chinese people, who regard it as a new form of socialist democracy. Although for most scholars, cyber democracy does not necessarily challenge the prudence of representative democracy, some intellectuals consider it to be an alternative to representative democracy.

Deliberative Democracy/Consultative Democracy

In today's China, various ideas of democracy, including direct democracy, autonomous democracy, participatory democracy, and cyber democracy, are becoming increasingly influential compared with liberal democracy. In addition, deliberative democracy has been regarded as a possible model for China's democracy by some officials and scholars. The concept of deliberative democracy was created in Western society. Ironically, deliberative democracy serves as a supplementary mechanism accompanying representative democracy in the West, but it is considered by some Chinese scholars and government officials as an alternative to electoral competition.

The popularity of deliberative democracy in China is partly attributable to the Chinese advocacy of political consultations between the ruling party and other political actors, which bears a resemblance to the idea of Western deliberative democracy. According to the system of multiparty cooperation and political consultation under the leadership of the CPC, the CPPCC has been established at both central and local levels. Although the Chinese practice of political consultation has a different implication from the Western idea of democratic deliberation, in Chinese, the term *consultation* (*xieshang*) shares the same characters as *deliberation*. Consequently, *consultative democracy* in Chinese (*xieshang minzhu*) becomes an equivalent for *deliberative democracy* in English, thus blurring the nuance between the two key words. Within this linguistic context, some Chinese scholars further argue that Western democracy is traditionally characterized by electoral democracy, whereas Chinese democracy is traditionally characterized by deliberative democracy (actually consultative democracy) since China has already established the CPPCC system, which symbolizes the characteristics of Chinese democracy or the Chinese model of democracy.

Other scholars assert that deliberative democracy and liberal democracy can support each other. This perspective is endorsed by the White Paper on China's Party System, issued by the Chinese government in 2007, which recognizes the importance of both electoral democracy and deliberative democracy. Furthermore, some scholars consider the two main Chinese political institutions—people's congress and people's political consultative conference—as examples of electoral democracy and deliberative democracy, respectively.

It is worth pointing out the significant difference between political consultation in China's context and deliberative democracy in Western backgrounds. However, there are also eye-catching experiments on the very sense of *deliberative democracy* in a few Chinese grassroots communities, such as the democratic discussion meetings in Wenling, Zhejiang Province. Other forms of community discussion councils and public forums have been created in a few localities thereafter.

Inner-Party Democracy

Compared with other proposed models, inner-party democracy is a widely recognized scheme for China's potential democratic model. Some scholars believe that China should first develop inner-party democracy since the CPC is the only ruling party normatively and empirically. Inner-party democracy is a legitimate tenet that has been accepted by the ruling party from the very beginning, and is characterized by criticism and self-criticism and a democratic lifestyle within the party. Since the 1990s, it has been reinvented by Chinese scholars as a strategy for China's democratization, while searching to democratize in the existing political system itself with one-party leadership. The logical sequence for political democratization in China, accordingly, is from the CPC to the outside of the ruling party, from the elite to the masses, and from intrasystem to extrasystem domains. This strategy was acknowledged officially when the party leader Zemin proclaimed to "promote people's democracy actively through the development of inner-party democracy" in 2001 at the 80th CPC anniversary gathering.

For some scholars, the basic principle for inner-party democracy is to handle correctly the relationships among the CPC's National Congress, the Central Committee and its Politburo, and Politburo Standing Committee in order to allow the party's National Congress and the Central Committee to perform their basic roles. The key approach to achieve this goal is to improve the electoral function of the National Congress and Central Committee by encouraging multicandidate elections (*cha'e xuanju*), which will create an institutionalized channel of interest aggregation for the party elites. These proposals have been more or less endorsed by the party since the beginning of the 21st century. Major measures being adopted or that are under consideration include improving the function of party congresses, expanding power sharing among party committee members, increasing information sharing among party members, developing a division of power within the Party, and allowing limited inner-party electoral competition. The CPC first introduced multicandidate elections in the 13th Party Congress in 1987 and expanded electoral competition in the 17th Party Congress in 2007. Such a reform program could possibly become the potential germination point of democratic transition and competitive elections in China.

Inner-party democracy, therefore, does not challenge the wisdom of representative democracy and electoral competition; rather, it tries to accommodate the plurality of social interests and the spirit of free competition by continuously expanding the scope of democratic elections within the party and finding a way to reconcile the model of "one-party dominance" with the model of "one-party pluralism" derived from the experiences of East Asian political development. In other words, supporters of inner-party democracy believe that in searching for a democratic model with Chinese characteristics, they can take a third path—an institutional compromise between single-party authoritarianism and multiparty democracy. In the opinion of some scholars, the strategy of building inner-party democracy or following the model of one dominant party with pluralism within, from a long-term perspective, may result in a chain reaction in the Chinese democratization process and would be an alternative to a multiparty system.

Democracy in Controversies: Liberal Democracy and Beyond

Though there is no lack of advocates of liberal (Western style) democracy in Chinese intellectual circles, more and more criticisms of liberal democracy have emerged in contemporary China, which has even been conceived as a wave of antidemocratic ideology in a general sense. Debates on democratic theories, in general, and liberal democracy, in particular, concern whether democracy is a good thing for China and whether China can find an alternative (such as deliberative democracy) to liberal democracy featured by electoral politics.

Democracy: A Good or a Bad Thing?

Supporters of democracy believe that democracy is a good thing for China. Some scholars contend that democracy is a universal truth, observing that a string of antidemocratic ideas, combined with the old thesis of "China uniqueness," is gradually developing and has severely blocked China's movement toward a socialist democracy. One of the influential arguments is that only democratic socialism can save China, and only democratic constitutionalism can fundamentally resolve the problem of political corruption within the government. Some scholars are more cautious in supporting democracy, proclaiming that democracy is neither a bad nor a good thing, but then is also the worst form of government except for all those other forms that have been tried, as Winston Churchill once said, or proclaiming that what China needs is a good democracy, not a bad one, paying more attention to substantial democracy rather than procedural democracy.

Opponents of democracy, by contrast, believe that democracy will do harm to China. Some argue that political order and social harmony are more important than other values and what China needs is not democracy but a limited government of checks and balances with the rule of law. Presupposing grassroots mobilization, popular elections, majority rule, and power politics, democracy would result in the "tyranny of the majority" in China's context. Some even bluntly declare that democracy is a bad thing and that democratization is the worst choice and would cause social disturbance and disaster in China rather than resolving the problems of political corruption and social inequality. Some scholars also dismiss the idea that democracy is a universal truth that can avoid the worst features of political life.

Arguments Against Liberal Democracy

Debates on general democratic theories in China focus on the value of liberal democracy, the mainstream of Western democratic theories. Liberal democracy is not the mainstream in Chinese academic discourse. On the contrary, various criticisms of liberal democracy and electoral competition are popular, particularly in official media. For some Chinese new-Leftists, furthermore, the ideal type of government is demarchy or lottocracy, a political system run by randomly selected decision makers who have been selected by lot. They argue that economic equality (rather than freedom) is the precondition for political equality, and representative democracy through the mechanism of free elections cannot guarantee an equal say for all citizens in public affairs. Unsatisfied with Western democracy, one scholar argues that once democracy is defined by any other concept, such as liberal democracy or representative democracy, it will lose its true meaning and value. The core value of pure or direct democracy is equality; hence democracy can be better guaranteed through the mechanism of selecting leaders and making policy by casting lots, in addition to enhancing political consultation, cyber participation, and working-place democracy. One can find a link between such an idea and Athenian democratic practice. A common thread running through these models is a strong commitment to egalitarianism, even at the cost of liberalism.

From the perspective of liberal democracy, open and fair elections are the bottom line of democracy, though it may be insufficient or unachievable without the support of other favorable conditions. In the Chinese context, however, many perceive competitive election as either a bad thing or as being irrelevant for democratic development. The new Leftists are not satisfied with liberal democracy because they want people's rule or direct democracy; meanwhile, the conservatives argue that competitive elections will bring about political chaos, money politics, social discord, and so on. For all the opponents of liberal democracy, the best solution for China's dilemma regarding democracy is to develop a democracy without elections.

Debates on Deliberative Democracy

Contrary to liberal democracy, deliberative democracy has gained more popularity in China than in Western democracies, and deliberative democracy in China is usually linked to multi-party cooperation and political consultation under the leadership of the CPC. As some scholars point out, this is a misreading or misunderstanding of deliberative democracy in its original meaning. As for the scholars who know well the theory of deliberative democracy, some forms of deliberative democracy, such as democratic discussion forums practiced in local politics and public hearings, are also heralded by them. In their opinion, institutional deliberation, as one form of democracy, can improve the quality of China's local politics and bring China to democracy. Some scholars expand the connotation of deliberative democracy to include other elements of democracy, such as elections, checks and balances of powers, rational discussion, and direct participation, and advocate the promotion of deliberative democracy as a rational strategic choice for China in developing socialist democracy. Some even suggest considering deliberative democracy as an alternative to electoral competition.

Other scholars, however, take a more cautious approach concerning the feasibility of deliberative democracy in China. They argue that there is a gap between China's political consultative system void of competitive elections and Western deliberative democracy based on representative politics. As an ideal and nonmainstream theory in the West, deliberative democracy cannot function as an alternative to representative democracy. The strength of deliberative democracy lies in its encouragement of rational thinking, dialogue, and participation, but it consumes more time and money than periodical elections. They argue, therefore, that the feasible step for China to take is to develop representative democracy based on elections, and without the context of competitive elections, true deliberation on an equal basis cannot take place, even at the grassroots level.

Divergence in the Chinese Model of Democracy

While the Chinese authorities declare that they intend to build a socialist democracy with Chinese characteristics and some scholars are keen to illustrate the Chinese model of democracy along with China's economic rise, many intellectuals in today's China believe that democracy is a universal concept and that there is no particular model for China's democracy. Democracy as practiced in China and other countries in the world should share some fundamental principles and institutional arrangements.

Some scholars believe, by contrast, that China has created a unique model of democracy, which is more advanced than that practiced in the West. As one scholar argues, China has experienced the most successful industrialization in the world, guaranteed people's rights and created a dynamic society through orderly political reforms, and enjoyed a leap forward because of the concentration of national resources. This suggests that the Chinese democratic model is the best one in the world. This perspective has been endorsed by the White Paper on the Building of Political Democracy in China, issued by the Chinese government in 2005. According to this white paper, there is no single, absolute, and universally applicable democratic model in the world.

Rather than highlighting the Chinese characteristics of democracy or the best democratic model in China, some scholars argue that China has not crossed the threshold of democracy at all but remains an authoritarian regime or posttotalitarian state. While Chinese elites have made great efforts since the reforms to promote grassroots democracy from below and inner-party democracy from inside, no substantial progress in terms of democratization has been made in the view of the liberals. Up to now, there is no consensus on China's democracy and its model. Philosophically, however, China's democracy will be a combination of universalism and particularism, which is accepted by most intellectuals.

Wei Hu and Gang Lin
Shanghai Jiaotong University
Shanghai, China

See also Communist Systems; Democracy, Direct; Democracy, Theories of; Democracy, Types of; Election by Lot

Further Readings

Fang Ning. (2009). Chinese democracy is the best one. *Hongqi Wengao* (Beijing), *2*, 4–9.
He Baogang. (2005). Institutions of deliberative democracy in China. *Journal of Zhejiang University* (Hangzhou), *3*, 13–21.
Hu Wei. (1999). Inner-party democracy and political development: Exploiting intra-system resource for China's democratization. *Fudan Journal* (Shanghai), *1*, 1–11.
Li Junru. (2005, September 27). What kind of democracy can China practice? *Beijing Daily,* p. 17.
Mao Tse-tung. (1975). On the people's democratic dictatorship. In *Selected works of Mao Tse-tung* (Vol. 4, p. 419). Beijing, China: People's Publishing House. (Original work published 1949)
Pan Wei. (2003). *The rule of law and the myth of democracy.* Hong Kong SAR: Hong Kong Social Sciences Press.
Wang Shaoguang. (2008). *Four lectures on democracy.* Beijing, China: Shenghuo Dushu Xinzhi Press.
Xie Tao. (2007). The model of democratic socialism and China's future. *Yanhuang Chunqiu* (Beijing), *2*, 1–8.
Xu Yong. (1997). *Villagers' autonomy in rural China.* Wuhan, Hubei, China: Central China Normal University Press.
Yu Keping. (2006). *Democracy is a good thing.* Beijing, China: Social Sciences Academic Press.

DEMOCRACY: MIDDLE EAST PERSPECTIVES

Notwithstanding the political change that has taken place to varying degrees during the past 2 decades in the countries of the Middle East and North Africa (MENA), an extensive region comprising the 22 members of the Arab League as well as Israel, Iran, and Turkey, Middle East scholars concur in the assessment that, with few exceptions proving the rule, MENA is still a democratic wasteland. Although most of the region's regimes run some sort of representative body or have a formally democratic political system, this statement holds true even when applying a minimalist Schumpeterian definition of democracy focusing on the legitimization of governance through regular free and fair elections. The relative absence of both democratic regimes *and* substantial democratization processes, especially in the Arab region, has led to intensive discussions about why the region has largely been unaffected by the "Third Wave of Democratization" (Samuel P. Huntington). To denote the region's alleged uniqueness in autocracy, economic misery, and the abundance of conflicts, the term *Middle Eastern Exceptionalism* has gained wide usage, even though many Middle East scholars feel uneasy about the term itself and its implications. While the assumption that MENA is the most conflict-prone world region and that conflict dynamics there follow a deviant logic as compared with non-MENA conflicts has been invalidated, many scholars also challenge the core assumption of Middle Eastern Exceptionalism—that is, that MENA proves to be exceptionally hostile to democratization. First, it is argued that, given the sobering record of persistent authoritarianism and stalled or aborted democratization processes in a considerable number of countries in Central Asia, Africa, and the former Soviet Union, MENA can hardly be referred to as being unique. Second, promising democratization processes in Turkey, a candidate country for membership in the European Union (EU), and Israel's consolidated democracy not only question the term's spatial scope but also reveal its lack of explanatory power when discussing global trends.

In order to approach the prospects for Middle Eastern democracy, this entry first gives a review of the obstacles to democracy in MENA as discussed in the academic literature. It then outlines recent sociopolitical developments, including the role of Islamist parties in the domestic political process and their impact on democratization.

Roadblocks to Democracy

The quest for the reasons that account for the absence of democracy in MENA has spawned an immense body of literature over the past years. At least five major determinants are put forward to explain why authoritarianism is still the dominant type of regime in the region: weak civil societies, patrimonialism, rentier structures, geostrategic location, and Islam. None of these determinants is without controversy, though. Before elaborating on the single determinant, it is worth noting, however, that the shift from the macrolevel of analysis in democratization theory to agency-based approaches as introduced by the seminal work *Transitions From Authoritarian Rule* by Guillermo O'Donnell and Philippe C. Schmitter (1986) was accompanied by the conceptual break with the "precondition" tradition. That is, theories of democratization that preceded agency-based transition theory took for granted that the establishment of democracy would be a by-product of modernization characterized by greater wealth, the formation of a bourgeoisie, more tolerant civic cultures, and the overcoming of economic dependency. But as democratic transitions in former authoritarian countries in Latin America and Eastern Europe have shown, that reality somewhat belied theory since those countries democratized without meeting the preconditions or prerequisites of democracy in advance. Thus, recent democratization theory argues that favorable socioeconomic conditions account for the survivability of democracy rather than the transition to it. It is, therefore, valuable that Eva Bellin (2004) reminds us of the pitfalls in the precondition paradigm, which seems to have been extended in the case of Middle Eastern studies, even though it is ill-founded from a theoretical perspective.

The following reasons are listed to illustrate the failure of Middle Eastern democratization: First, civil society, which is commonly assumed to play an essential role in the development of a vibrant democracy, is overwhelmingly judged to be weak in MENA. However, assessments on civil societal structures in MENA are ambivalent. While some authors even wonder whether "there is any civil society at all in the Arab Middle East," others like Saad Eddin Ibrahim are far more confident, detecting a potential for grassroots pressure in Middle Eastern civil societal organizations. To be sure, a balanced evaluation of the strength of civil society depends on the differences among individual MENA countries. Concerning Egypt, Syria, and Lebanon, for instance, the Arab Human Development Report 2005 positively speaks of an increasing role of civil societal organizations in the political public sphere. On the other hand, some authors dislike the Islamist domination of the civil society sector, a fact that need not necessarily appear in a bad light since, even among Islamists, there is broad agreement that enlarging a democratically engaged public sphere is top priority for MENA development.

Related to this argument is the assumption that a patriarchal tradition of power is deep-rooted in the Arab-Islamic culture and thus impedes democratization. The persistence of patrimonial regimes, however, is more convincingly elucidated by the rentier state paradigm than by mentality. As Giacomo Luciani has noted, rentier states, that is, states whose revenue is predominantly derived from nonproductive sources of income, such as oil or other foreign sources, and whose expenditure constitutes a substantial share of gross domestic product (GDP), are less dependent on the extraction of wealth from their populations to finance the state. Accordingly, there is no need for accountability on the part of the ruling elites whose actions remain unchecked for the reason that, from an alleged population's point of view, it is not imperative to stand up to arbitrary taxation and demand greater political participation ("No representation without taxation"). Political quietism is furthermore evoked by governmental distribution of rents, which, in addition, solidifies the existing patterns of political patronage. The share of the private sector in GDP remains marginal and accounts for both the calamitous economic performance of rentier states and the impediment to the emergence of a large middle class, which is commonly regarded as the germ cell of democracy.

The "oil factor" leads to another major roadblock for Middle Eastern democracy: the geostrategic location of MENA states. It is contended that the desire of the United States and Europe for stable Arab regimes outweighs their desire for democratization, and thus potentially destabilization, of the region since only stable regimes can safeguard Western energy supply and contain the terrorist threat. Western claims for a democratic Middle East are hence undermined by Western economic and military aid to MENA regimes, in the eyes of local human rights activists. The geostrategic determinant, furthermore, subsumes the impact of the Arab–Israeli conflict that is often believed to inhibit democratization as the regional authoritarian regimes argue that the external threat of Israel renders the maintenance of a strong military and security apparatus necessary, which, in fact, turns out to be a coercive apparatus for internal repressive measures in the first instance, as these regimes use the Israel threat as an excuse for internal repression.

Finally, Islam is probably the most contentious point of debate. Some well-known scholars maintain that Islam is utterly inconsistent with Western notions of democracy and pluralism. By contrast, there exists a voluminous literature demonstrating that Islam and democracy are not only compatible but can also be complementary. Here, it is necessary to distinguish between two levels of analysis. First, it is important not to confuse Islam with Islamism, with the latter being defined here and commonly as the active assertion and promotion of beliefs, prescriptions, or policies that are supposed to be Islamic in character—a definition that, by its nature, awkwardly encompasses *jihadi* groups, such as Al Qaeda, as well as the Syrian liberal reformer Muhammad Shahrur. The Islamic authoritative scriptures (Qur'an and Hadith) leave wide margins for interpretation, especially when it comes to the political, social, and economic relations among humans as well as the relations between humans and governmental organizations. Having said that and bearing in mind the variety of Islamic doctrinal schools and traditions, one can understand the existence of a wide spectrum of conflictive interpretations by Islamic scholars (ulema) or Islamist activists, ranging from the liberal interpretations of, for example, the Egyptian Islamist *Wasat* Party to the fundamentalist tenets of the influential Egyptian Islamic scholar Yusuf al-Qaradawi. In addition, empirical findings have shown that the support for democracy as an ideal is high among Muslims and, more important in this context, that Islam has less influence on political attitudes than is frequently presumed (Steven Hofman, 2004).

This list of reasons, which are supposed to account for the absence of democracy in MENA, is not comprehensive. Other causes associated with the major categories introduced above are also highlighted, such as the low level of education, the weakness of national identities based on artificial Middle Eastern state formation, or the remoteness of a democratic neighborhood. While various reasons are commonly cited in a cumulative way to dwell on the Middle Eastern failure of achieving democracy, few authors contest that the region has undergone important sociopolitical changes in the past few decades. Rather, it is now the scope, depth, and intention of political reforms that are debated.

Recent Trends in Sociopolitical Development

Over the past few years, the MENA countries have launched a number of reforms in response to domestic and global challenges. The dimensions of political, economic, and social crises that the Arab Middle East faces are vividly captured in the recently published Arab Human Development

Reports. These reports, indeed, demonstrate that Arab regimes lag behind with regard to some significant political and economic reforms. However, the steps that have been taken in several MENA countries to countervail the present crises may also be considered as a basis for notable change in the medium and longer term. Subsequently, the ambivalent record of reforms in MENA again opens the door for scholarly discussion.

Regarding Middle Eastern economies, the provision of finance by the International Monetary Fund and the World Bank has, in return, led to structural reforms in key Arab states, such as Tunisia, Egypt, Jordan, and Morocco. These reforms induced programs of economic liberalization that opened up recipient economies to the world economy. This so-called *infitah* policy has raised hope among proponents of modernization theory that economic liberalization will be followed by political openings in the long run. In terms of political reform, the MENA region has witnessed slight improvements in political rights and civil liberties. The Freedom House Survey 2009 asserts that during the Bush presidency, nine of the region's countries experienced some advancement on the Freedom in the World scale, including several of the Gulf States and Saudi Arabia. Yet despite the first formally competitive presidential elections in Algeria (2004) and Egypt (2005), the first Saudi municipal elections in 2005, the introduction of a new family code in Morocco in 2004, and the appearance of political protest movements (e.g., *Kifaya* in Egypt and the Damascus Declaration in Syria), all of which have been interpreted as important steps toward more political liberalization, no major breakthroughs had been detected in MENA by 2009. Israel is still the only country in the region that the Freedom in the World 2009 survey (by Freedom House) classifies as *free*, whereas only six MENA states are categorized as *partly free* (Turkey, Jordan, Morocco, Lebanon, Bahrain, and Kuwait). The rest fall into the category of *not free*.

However, events in 2011 might turn the political landscape of the Arab world inside out. On January 14, the long-serving Tunisian President Zayn Al-ʻAbidin Bin ʻAli surrendered to public demands and, following a relatively short period of massive popular protest, fled the country. The seemingly effortless fall of Bin ʻAli, whose security apparatus was deemed one of the strongest and most oppressive in the region, came to be known as the Jasmine Revolution and incited further popular upheaval across the Arab world, primarily borne and engineered by the region's urban, marginalized, and well-educated youth. Only a short time after, on February 11, Egypt's president, Hosni Mubarak, was swept away. While other Arab leaders may still meet the same fate, it remains to be seen whether the so-called Peaceful Youth Revolution (*thawra al-shabab al-silmiyya*) in the Middle East will eventually result in liberal democracies or not.

Skeptics admit that today a rapid and extensive political transformation process is underway in MENA but point to the robustness of Middle Eastern authoritarianism. In light of the ambivalent inventory of reforms predating the Arab uprising, only few authors saw the Arab world as being on the path toward democracy. Rather, there was widespread agreement that the region's authoritarian regimes had learned not only to accommodate themselves to new political, economic, and social conditions but also to shape them according to their own best interests. Steven Heydemann hence speaks of "Authoritarian Upgrading" and illustrates how Arab regimes have succeeded in managing political contestation, benefiting from selective economic reforms, controlling new communication technologies, and stemming as well as appropriating civil society. Or, in other words, they have up to now skillfully handled Western democracy promotion by means of superficial reforms while simultaneously retaining domestic authority.

The Rise of Islamist Competitors

Yet another trend characterizes the recent sociopolitical development in the region—the rise of Islamist competitors to authoritarian rule. Within the past 2 decades, Islamist movements (*haraka*) such as the Muslim Brotherhood (*al-ikhwan al-muslimun*) have become popular grassroots organizations with their associated political parties (*hizb*) being capable of winning the majority of votes in most MENA countries, if free and fair elections were held. From Morocco to Iraq, Islamists are represented in the national parliaments, if granted access to the political system, or, like the *Mouvement pour la Société de la Paix* (MSP) in Algeria, even participate in government coalitions. This thriving of Islamist parties is closely observed by Western academics and politicians alike. Still, many voices disapprove of and fear the Islamists' political successes. In the meantime, however, the majority of Middle East scholars distinguish between moderate and radical Islamists. While the latter aim at the creation of an Islamic state through revolution in the Qutbian tradition (or seek to maintain it, as is the case, for instance, with the Shiite ulema appointed to the *Council of Guardians* in the formally democratic Islamic Republic of Iran), the former are prepared to pursue their political agenda by peaceful means within the existing political institutions. This is a more balanced approach, which gives consideration to Middle Eastern realities. It is far from reasonable to look at Turkey's ruling Islamist Justice and Development Party (*Adalet ve Kalkınma Partisi*: AKP) in the same way as at the fundamentalist *Hizb al-Tahrir*, which is banned in many MENA countries. Vali Nasr, for instance, counts the AKP among "Muslim democrats." According to Nasr, Muslim democratic parties, which are found mainly in the non-Arab Muslim world, are akin to Christian democratic parties due to their commitment to democratic values, their moderate religious ideology, and their systemic modus operandi. Such moderate Islamist parties, however, are also emerging in the Arab Middle East. Or, more precisely, as Islamist parties are overwhelmingly composed of radical and moderate wings, moderates there are getting the upper hand over their more radical party members.

The success of Islamist parties must not only be attributed to the eroding legitimacy of secularist and nationalist elites in MENA but also to the ability of their Islamist mother organizations to synthesize the interests of the lower and the middle classes. Besides, Islamist movements are popular because of their welfare activities, which provide broad social strata with basic health, nutrition, and educational services, a task insufficiently performed by public institutions. Also, as a result of anticipated electoral successes on account of a strong popular backing, Islamist parties have turned toward political participation within existing state institutions—all the more as former strategies of overthrowing regimes using violence have failed. This change of strategy can, in addition, be substantiated by the rationale of rent seeking: The middle classes who are incorporated into Islamist movements to a large extent do not engage in conflicts with the state but rather hope for more upward mobility and access to rents from within the state apparatus. Evidently, the inclusion of Islamists in the political process also presupposes the preparedness of the regime to include rather than exclude Islamists.

At present, Islamist parties are on a learning curve, and the moderation of many Islamist parties arises from their inclusion in the political process. After their moving into parliament, Islamists have to assume responsibility for taking action as well as refraining from doing so. Like every political party, Islamists have to compile party platforms, which give answers to domestic and global challenges and justify their positions in public. Once accountable to the electorate, they swiftly realize that there are no such easy remedies for urgent social and socioeconomic problems, such as just proclaiming "Islam is the Solution" (*al-Islam huwa al-hall*). The day-to-day work in parliaments or in parliamentary committees also leads to an institutionalization and modernization of party structures. Similarly, it requires getting along with former ideological opponents. To convince non-Islamist parliamentarians, Islamists learn to argue on a political and economical basis in different policy fields instead of incessantly resorting to religious reasoning. All this leads to the professionalization and de-radicalization of Islamist parties, as can be observed in the cases of the Moroccan Justice and Development Party (*Parti de la Justice et du Développement*, PJD) or the Algerian Movement for the Society of Peace (MSP). Within the party leadership, new technocratic elites are eclipsing the old religious elites step by step, a process that involves both the gradual establishment of inner-party democracy, including internal party strife as a matter of course, and growing independence from their mother organizations.

Obviously, not all Islamist parties chose to take the path of political participation or, as mentioned before, were barred by the regime from doing so. However, all Islamist parties that did embark on the strategy of participation have been faced by what is best described as the "participation dilemma." That is to say, as Islamist parties are co-opted into the regime, they run the risk of losing credibility in the eyes of their supporters. Votes of Islamist parties in Jordan, Algeria, Morocco, and Yemen demonstrate that Islamist parties, accordingly, have performed badly in elections once they entered parliament or the government—all the more as the hoped for exertion of influence on political decision making often proved to be unrealistic. Since the political system that they presently operate in does not allow for serious policy changes, Islamists often confine themselves to playing the role of moral watchdogs in parliament. Nevertheless, in this way, Islamists have become a major pillar of the regimes since by their very participation in the political process they bestow legitimacy on the regimes and foreclose the emergence of a more powerful opposition.

Moderate Islamists: Partners of the West?

Even though there is no doubt that there have been significant changes in the ideologies of many Islamist parties and their course of action, the question of whether this change is genuine and can be trusted is a highly controversial one. Many hold that even moderate Islamists will only show their true colors after having assumed power. While it is virtually impossible to disprove the latter assumption *ex ante*, drawing a balance sheet of academic findings about the political performance of moderate Islamists in the past can be helpful. Whether out of strategic considerations or conviction, moderate Islamist parties have stipulated democracy in their party platforms. Here, democracy is no longer understood as a Western product, rather general procedural aspects are emphasized. Islamists strongly advocate transparency as well as the establishment of and abidance by democratic standards. By way of inclusion into the parliamentary system, Islamists have learnt to focus on national interests and to put ideology on hold. In this way, Islamists are guided by increasing pragmatism when attempting to handle political challenges. Furthermore, Islamists have shown their willingness to seriously work in coalition with secular parties and, as is the case with the Yemeni Joint Meeting Parties, also with other denominational forces. Having said that, to this day Islamist parties have not yet completed the process of moderation, and it is still uncertain whether their interpretations of Islamic law (*shari´a*) are flexible enough to fully reconcile religious with democratic values.

Islamist parties have taken a firm stand, too, with regard to another important policy field, the economy. An analysis of party platforms shows that Islamists favor a social market economy modeled on continental European examples. That is, they clearly reject a state-directed economy while approving a state-led regulation of the public sector. And they demand the strengthening of the private sector while endorsing public welfare spending. Hence, Islamist economic programs are directed to their party clientele, the middle class. And it is, ironically, the Islamist parties that are committed to agreements with regard to the Washington Consensus.

Against the background of the dismal prospects for democracy in MENA, on the one hand, and the rising relevance of Islamist parties, on the other, more and more Middle East scholars contend that democratization in MENA is impossible to reach without the participation of Islamist parties. Thus, Western governmental elites and scholars alike have begun to address the issue of whether or not moderate Islamist parties can be seen as reliable partners in future democratic developments.

Rachid Ouaissa and Jens Heibach
Philipps-Universität Marburg
Marburg, Germany

See also Authoritarian Regimes; Christian Democratic Parties; Civil Society; Clientelism; Conflicts; Democracy, Theories of; Democratization; Fundamentalist Movements, Islamic; Islam; Islamist Movements; Neo-Patrimonialism; Opposition; Participation; Parties; Pluralism; Political Culture; Religious Movements; Social Movements; Terrorist Groups

Further Readings

Bellin, E. (2004). The robustness of authoritarianism in the Middle East: Exceptionalism in comparative perspective. *Comparative Politics, 2,* 139–157.

Browers, M. L. (2006). *Democracy and civil society in Arab political thought: Transcultural possibilities*. Syracuse, NY: Syracuse University Press.

Diamond, L., & Plattner, M. F., & Brumberg, D. (Eds.). (2003). *Islam and democracy in the Middle East.* Baltimore: Johns Hopkins University Press.

Freedom House. (2010). *Freedom in the world.* Washington, DC: Author. Retrieved August 3, 2010, from http://www.freedomhouse.org/template.cfm ?page=15

Hofman, S. R. (2004). Islam and democracy: Micro-level indications of compatibility. *Comparative Political Studies, 6,* 652–676.

Khatab, S., & Bouma, G. D. (2007). *Democracy in Islam*. London: Routledge.

Luciani, G. (Ed.). (1990). *The Arab state*. London: Routledge.

Nasr, V. (2005). The rise of "Muslim democracy." *Journal of Democracy, 2,* 13–27.
Salamé, G. (Ed.). (1994). *Democracy without democrats: The renewal of politics in the Muslim World.* New York: I. B. Tauris.

Democracy: Russian Perspectives

The first use of the word *демократия* (*demokratiia*) in Russian can be traced to the translation of Samuel Pufendorf's (1718) *Introduction to the History of the Principal Kingdoms and States of Europe*. But this does not mean that there was no idea of democracy in Russia before that time. Russian chroniclers or statesmen have, centuries earlier, described some democratic practices with words such as *самодержавие* (*samoderzhavie*—literally "*self-consolidation*"), *самовластие* (*samovlastie*—literally "*self-rule*"), and *самоволие* (*samovolie*—literally "*self-will*"), and they called some Cossack communities "a kind of a republic" using a transliterated Polish word *речьпосполита* (*rzeczpospolita*) or simply called democratic orders "free" (*свободный*—*svobodyi*) or "unrestricted" (*вольный*—*vol'nyi*). This entry discusses some of the historical roots and various interpretations of democracy in Russia.

Despite the relatively late use of the term, *democracy* was quickly given a very specific Russian implication. Democracy was considered to be deeply inherent in the everyday life of the Russian peasant community—*obschina*—and exceptionally specific in this regard, nothing like its plain Western counterpart. The 19th century was a time of profound social crisis and ongoing preparations for the liberation of serfs and subsequent modernizing reforms. It was the time when talk about the democratic instincts of Russian peasants and the archetypal centrality of pure and incorrupt (direct) democracy to Russian *obschina* were common in intellectual circles. In fact, democratic potential can be traced to primordial tribal times in any tradition. As for Russia, such democratic ways were fairly well expressed in history and were still alive in local communities in the 19th century and even later.

Beginning around the 9th century, the poleis of Eastern Slavic tribes and tribal federations interacted with the equally rudimentary military democracy of Varangian (*druzhina*—princely retinue, literally "camaraderie") and set up a common military–trade infrastructure. Their interface contributed to the development of a patrimonial power structure of the *Kievan Rus*' principality, where interactions between the prince and the people produced legitimizing effects. Polis structures with their democratic potential (such as *veche*—*ecclesia* and/or *boule* in various cities—literally *talk, deliberation, council*) were important elements of early democratic elements in Russia, as were *ves' grad* (general assembly or polis, as a partner to the prince, literally *all the city*), *startsy gradskie* (city elders), and *sbor* (assembly). Equally significant were conciliar structures such as *druzhina* (early princely council, literally *camaraderie*) and *duma* (late princely council, literally *thought*).

The Mongols (the "Golden Horde") invaded Russia during 1236 to 1242. With the conquest of Rus' by the Horde, princes emerged as the only authority to control political order in Russian dependencies of the Horde in the eastern and southern parts of the former Kievan Rus'. Their rule was granted by Khans (*iarlyk*—mandate, Mongolian *jarligh*—literally *order*). Actually, the Khan was the only source of all power. The power was delegated from the top down. Still, there was only one but significant exception. Princes also had to maintain the Christian *derzhava*l (literally *something holding together*), which united them with the people and domestically was an alternative source of power and princely authority.

In ancient and medieval Rus', ideas of democracy and autocracy were not opposed to each other but in fact closely interwoven. In fact, in medieval Russian texts, the terms mentioned earlier—*samoderzhavie* ("self-consolidation"), *samovlastie* ("self-rule"), or *samovolie* ("self-will")—were used to denote both democracy and autocracy. The difference was contextual. It depended on who was the self in question—the whole community or the ruling authority or, more typically, both, integrated into a single body politic. It was conceptual substance (consolidation of power) that was important, not the form of the consolidation—monocentric, polycentric, or dispersed.

The Russian word *samoderzhavie* is usually translated as "autocracy." The word actually emerged as a Slavic counterpart of the respective Greek term *autokratia*. But the translation was not exact. While the first components of both words (*auto-* and *samo-*) may be considered fairly equivalent as denomination of Self, the second ones were somewhat different. Both implied power but of a different kind. The Greek *kratos* denoted coercive and instrumental power, mingling the ideas of "military or physical supremacy" (i.e., *krat-*) and "hardiness" (i.e., *kartu-*) (Émile Benveniste, 1969). The Slavic *derzhava* connoted integrative power of holding people together originating from *dher-* ("to hold together"). So while *autocracy* suggested coercive power of an unrestricted Self over one's subordinates, *samoderzhavie* stood for the power of self-integration. The reason was that a fairly primitive tribal power structure would not differentiate between the power emanating from the tribal community and the power of its chief. Conceptually, it implied that both authorities and people are fully integrated into a single body politic. In fact, this conceptual scheme has been so strong that it produced one of the most widespread Soviet slogans "People and the Party are integral" (*narod i partiia ediny*—literally "people and the party are one, single, united"). In fact, any ruling authority in Russia has been obsessed with the idea of unity with the people. Typically, Putin's dominant party is called *Edinaia Rossia* (literally "single or integral Russia") and not *Obyedinennaia Rossia* or "United Russia" as it is inaccurately translated.

Authority Versus People

The emergence of the Grand Duchy of Muscovy in the late 15th century was a great act of emancipation from nearly 3 centuries of Mongol external domination, known as the *Tatar Yoke*. Around that time, Ivan III Vasilevich (Ivan the Great) of Muscovy and his brother-in-law Stefan cel Mare the Great of Moldova introduced a conceptual innovation. Claiming political independence from the Golden Horde and the Ottoman Empire, they called themselves *gosudar'* and *gospodar'* (two alternative forms of the same word), respectively, and called the polities they controlled their "households"—*gosudastvo* and *gospodarstvo*, respectively. In current Russian, *gosudastvo* is a standard term for the state. The word *gosudar'* is just a form of the Old Slavic word *Gospod'*—the Lord. Its meaning can be traced back to the Indo-European notion of authority over aliens. In other words, the notion referred to a very important function of distinguishing kin from alien and deciding which aliens could be treated as adopted kin and which were to be rejected as enemies.

To stress his exceptional status, Ivan III demanded to be called not just *gosudar'* but *velikii* (great) or *samoderzhavnyi* (independent, self-integrating). Typically another Russian polity, the Novgorod Republic, was often called *Gospodin Velikii Novgorod*, which can be translated as the Great Overlord Novgorod or more accurately Great Sovereign Novgorod. In any case, both Muscovite and Novgorod claims to the status of *samoderzhavnyi gosudar'* implied that both the household lord—princely or republican—and his people were integrated by the same *derzhava*, which made them all free. The Grand Duchy of Muscovy reemerged as a sovereign (*samoderzhavnaia*) Russia in its first revolution of self-integration, as Russia's emancipation from the outside threat to its very existence turned into subordination to the very authority that was considered instrumental in bringing about that emancipation. Thus, the grand act of total emancipation immediately turned into ultimate subordination. This pattern is being repeated in the later history of the country as well.

Development of the Russian polity was uneven and contradictory. Still, new important elements of a democratic nature appeared including a system of representative bodies. In 1549, Ivan IV (Ivan the Terrible) convoked the so-called *Sobor Primirenia* (Assembly of Reconciliation). It was composed of three bodies—*Boyar Duma* (Royal Council), *Osveschennyi Sobor* (Blessed or Holy Assembly), and *Zemski Sobor* (Assembly of the Land). A short-lived attempt was made to create a structure of four "estates"—tsar, nobility, clergy, and commoners, with the last three having their assemblies.

Since the mid-16th century, all kinds of assemblies were more or less regularly convoked. *Zemski Sobor* was an important but subordinate body under Ivan the Terrible. It was essentially used to outbalance *Boyar Duma*. *Sobors* of the late 16th and early 17th centuries were instruments of political manipulation rubber-stamping adventurous takeovers and conquests. *Zemski Sobor* of the "tsarless" period of 1610–1613 became an actual legislative and even executive authority. In 1613, it had carried out a series of regional *sobors* and a special joint session to elect Mikhail Romanov as a new tsar and the founder of a new dynasty. Under Mikhail, *sobors* met regularly until 1622; their meetings resumed in 1632, but their roles changed. They were only to consent to the tsar's decisions and ceased to be convoked after 1684.

The long and dramatic story of changing concepts and institutions of Russia (*gosudarstvo*, *tsarstvo*, *derzgava*, *otechestvo*, etc.) arrived at a critical point in the mid-17th century. In Europe, it was the time of termination of the epoch of confessionalism marked by the Peace of Westphalia (1648). European sovereigns were emancipated from a mandatory subordination to transcendental authority and could rely on *ratio status*. They had to rationalize relations with their subjects by developing networks of bureaucracies and representative bodies. In Russia, the time was equally decisive. It reemerged again as a sovereign, *samoderzhavnaia* power in its second revolution of self-integration after the Times of Trouble of the early 17th century, marked by famine, the Polish invasion, and the end of the Rurik dynasty. This time, it was the so-called Velikaia Russkaia Samoderzhavnaia Revolutsia (Great Russian Autocratic Revolution; see Yuri Pivovarov & Andrei Fursov, 2001), in which the new Romanov dynasty emerged following Polish domination. Again, radical emancipation ultimately became subordination as history repeated itself, and domination of subjects by their rulers continued. In contrast to Western Europe, which was territorially divided, Russia or rather Northern Eurasia was territorially integrated. In Europe, sovereign states developed agencies of various levels to circulate the power. In Russia, all former political agencies were reduced to insignificance, and the tsar alone emerged as the sole authority to his *tsarstvo* and his people. With all the centralizing effects of the Great Russian Authoritarian Revolution, a country as big as Russia simply could not be governed by using only one simple principle of integration of power and the populace, which was the main feature of the Russian system that ensued after the Great Russian Autocratic Revolution. Many other patterns of rule, management, and decision making were applied particularly at geographical and social peripheries and specific domains of activities. It is there that undeveloped or even primeval patterns of democratic governance survived, including local village communities or small townships typically called *sloboda* (literally "free settlement"), military orders of Cossacks (literally "free rambler" in Turkic), religious and ethnic self-governing communities, companies of explorers (important for Russia's frontier of unexplored land in Siberia, the steppes, and the Far North), merchants and artisans (so-called *artel'*), and so on.

It took nearly 2 centuries (and more than a century of Europeanization) to expand the polity to such an extent, and for practical reasons, a vast network of institutions had to be developed to regulate relations between individuals, groups of individuals, and the imperial superstructure. Along with democratic patterns of operation at the peripheries of all kinds, these institutions were influenced and distorted by the Russian system but were to a lesser or greater extent resistant to its sway.

During that period two important political notions emerged: (1) *Otechestvo*—Fatherland—a nation with a long conceptual history from collective heritage within lineage through *votchina* and (2) a member of the nation typically called *Syn Otechestva*—Son of the Fatherland. Those two notions were extremely important alternatives both to the power and to the populace. Highly personalized power was depersonalized, and impersonal populace turned into a community of personalized sons of the nation.

By the beginning of the 19th century, Russia was ready for another revolution. It was accelerated by the Napoleonic invasion. The Patriotic War of 1812, called in Russian *Otechestvennaia voina* (the war that sons of the Fatherland waged to save their Fatherland), turned into the third revolution of self-integration. Russia reemerged again as a sovereign, *samoderzhavnaia* power destined to maintain the political and moral order of post-Napoleonic Europe as Alexander I the Blessed solemnly declared at the Vienna Congress. Another grand act of total liberation from foreign occupation turned into strengthening the autocrat and his or her power.

Still it was not all that simple and one-sided. Endogenous political developments coupled with efforts to Westernize and integrate into European political order produced significant changes. Probably the most decisive factor was the emergence of a notion of people that would cover not only the populace but also a potential actor safeguarding Russia itself. Typically it was "silent," for example, in the symbolic scene of the election of the tsar in Pushkin's "Boris Godunov," but it was able to act spontaneously and decisively during crucial, "revolutionary" moments such as *Otechestvennaia voina*. An important reflection of the changes and of a new balance of essentials of Russian politics manifested itself with the declaration of the guiding ideological formula—*pravoslavie* (orthodoxy), *samoderzhavie* (autocracy but also sovereignty and self-integrity), and *narodnost'* (nationality or people mindedness, ability to think and act as the integral Russian people). It is true that the formula coined by the Russian minister of education Sergei Uvarov in 1833 was a counterpart of the formula of the French revolution—*Liberté, Égalité, Fraternité*—and to that extent was "the Russian version of a general European ideology of restoration and reaction" (Nicholas Riasanovsky, 2005, p. 133). But it is equally true that it redefined the power setup within the Russian System. *Pravoslavie* reaffirmed the old principle of sacral *derzhava* with all the power coming from above. *Samoderzhavie* provided pragmatic *derzhava* of the autocrat and his *Polizeistaat* administration and also for Russia's sovereignty and integrity. *Narodnost'* indicated that in practical terms Russian people were the core frame of reference for conceptualizing and enacting politics. Conceptually, it implied that the Russian people along with the Supreme Lord and the Autocrat were a source of power. The modern principle of organizing politics bottom-up was implicitly recognized. As we can see it now, it was a major precondition to anticipated modernization and still murky democratization.

The tripartite formula was a step forward to rationalize political integrity not only conceptually but also functionally and structurally. It was confronted with the strong determination of *Slavophiles*, such as Ivan Kireevskiy and Alexei Khomiakov, to work out a new holistic idea for Russia that would be essentially based on the integrity of Orthodox ecclesia and Russian *obschina*. They introduced the idea of *Sobornost'*—holistic spirituality integrating worldly communities of Orthodox Russians. This new idea was potentially extremely revolutionary since it could imply that there was no need either for absolutist *Polizeistaat* machinery or for the state as such. *Sobornost'* would rely not on the mundane practicalities of rule and subordination but rather on traditions of peasant and *boyar* conciliarism having clear connections with the *sobors* of the patriarchal Muscovite past and the romanticized spiritual integrity of the people and the tsar.

Sovereignty of the People

In the discourse of, first, Alexander Herzen and then Nikolay Chernyshevskiy and his associates, *samoderzhavie tzarei* (autocracy of the tsars) was opposed to *samoderzhavie naroda* (popular sovereignty). A clearly modern way of thinking was embedded into the national tradition. Integration was a clear priority but could be achieved only on the basis of *narod* (the people) becoming a true sovereign or *samoderzhets* (autocrats). Early Russian democrats were very optimistic about the prospects of peasant democracy and, hence, socialism in Russia. In 1862, they created a revolutionary organization *Zemlia i volia*—land and liberty, or rather unrestricted self-will (Russian *воля*, English *will*, and Greek *βουλή*, all derive from the *uel*—"to wish"). The name of the organization clearly demonstrated the objective to surpass the moderate results of the emancipation of serfs and to achieve a far more radical redistribution of land alongside with provision of individual liberties of all kinds. This venture was short-lived. It was brutally crushed a year later when the Polish uprising provoked harsh political repressions throughout the whole country.

Soon thereafter, in the late 1860s and 1870s, a broader movement of *narodniks* (usually translated as "populists," but a more exact rendering would be "people minded, concerned with people's cause") emerged. From 1873 to 1875, many narodniks left the cities for villages on a mission of *khozhdenie v narod* (literally "going into the people") in an attempt to reintegrate with the peasantry and thus bring about moral and social revolution. This campaign failed. The radicals among the *narodniks* joined terrorist organizations, such as *Narodnaya Volia* (People's Will), which was active from 1879 to 1883. Some others joined the anarchist movement, inspired by Mikhail Bakunin and Petr Kropotkin. More moderate and liberal-minded persons accepted a strategy of "small ventures" and became active in local self-government—the *Zemstvo* movement.

In practical terms, *Zemstvo*—a name for a local representative body, literally "community of a land")—was one of the most significant achievements of the so-called Great Reforms. In 1864, Alexander II created self-governing bodies in a number of provinces and localities. Their number as well as competence gradually increased. *Zemstva* ran educational and medical establishments, developed all kinds of infrastructural projects, and so on. What is particularly important is that they practiced elections to representative bodies, for example, representative councils (*zemskoie sobranie*) and controlled executive ones (e.g., executive boards—*zemskaia uprava*). This was an invaluable democratic experience for Russia. *Glasnost'* (usually translated as "openness" but deriving from the Russian word for "voice" thus meaning "voicefulness") became an important key word. It implied accountability of authority and importance of public opinion in the running of reforms (Bruce Lincoln, 1981). *Perestroika* ("rebuilding") was another key word of the time.

With all the validity of local and autonomous self-government of specific communities, the imperative of *Volonté Générale* of the entire people still prevailed. Despite the growing diversity of the political and intellectual landscape of postreform Russia, popular feelings and intellectual efforts focused on ideas of political integrity. One of the most significant instances was a notion of *samoderzhavnaya respublika* (autocratic republic) introduced by the Founding Father of Russian liberalism, Konstantin Kavelin (1818–1885). His point of departure was "organic integrity ("oneness") of the authority and the people." On that basis, he concluded, "Since the people (collective singular) no doubt in its very essence are (literally *is*) autocratic, so single with it authority, eo ipso must be autocratic." He further developed his prognosis:

> Tsar is the only and most secure bulwark of peasantry against aristocratic and citizen [*meschanskikh*, literary "town dwellers" meaning "bourgeois"] constitutions. In future as well he is the best security against emergence of any privileged ruling classes. There is no doubt that by all its mass Russia would follow only the autocratic, i.e. free Tsar, who is independent either of boyars or of plutocrats. The history itself makes us create a new, unprecedented and unique political order that no other wording would fit but *autocratic republic*. (Kavelin, 1989, p. 436 ff).

A very similar political order emerged after the October Revolution of 1917. But the project, as well as the conceptualization, was different. During the revolution in 1905, Lenin started to speak about Soviet Power as a new form of political organization or direct democracy of the masses. In 1917, the slogan "All power to the soviets!" (*Vsya vlast sovyetam!*) was used by the Bolsheviks to successfully crush Kerenski's Provisional Government and to establish the new political order.

Democracy of Soviets

Both sovereign and autocratic (*samoderzhavnaia*) power reemerged again in its fourth liberation revolution of self-integration. This time it was the so-called *Velikaia Oktiabrskaia Sotsialisticheskaia Revolutsia* (Great October Socialist Revolution). Again, the grand act of total emancipation turned into outright dictatorship. Soviet Power was highly ambivalent: On the one hand, it relied on mass participation and thus had a clear democratic calling coupled with the institutional form of direct democracy of Soviets—"Councils of Workers and Peasants Deputies"—but on the other, the emergent system could only be run by a highly integrated and disciplined—a new type of—vanguard party. Lenin, in his seminal 1920 book, *"Left-Wing" Communism: An Infantile Disorder*, clearly fixed a hierarchy of power: leaders, party, (working) class, and masses.

Democratic centralism was an answer to the practical running of the country. Initially, it was a set of principles of internal organization of the nascent Russian social-democratic party advocated not only by Bolsheviks but also by Lenin personally. In its fuller version, democratic centralism implied five key points:

1. election of all party organs from bottom to top and systematic renewal of their composition, if needed;

2. responsibility of party structures to both lower and upper structures;
3. strict and conscious discipline in the party—that is, the minority must follow the majority decisions until such time as the policy is changed;
4. decisions of upper structures mandatory for the lower structures; and
5. cooperation of all party organs in a collective manner at all times and, correspondingly, personal responsibility of party members for the assignments given to them and for the assignments they themselves create.

With the consolidation of Soviet Russia, those five principles were introduced into the internal machinery of the Soviet system of rule with very minor adjustments of a purely technical nature. Democratic centralism was formally fixed in Article 7 of the Soviet Constitution of 1977 as a principle for organizing the state:

> The Soviet state is organized and functions on the principle of democratic centralism, namely the electiveness of all bodies of state authority from the lowest to the highest, their accountability to the people, and the obligation of lower bodies to observe the decisions of higher ones. Democratic centralism combines central leadership with local initiative and creative activity and with the responsibility of each state body and official for the work entrusted to them.

Another way to describe the Soviet rule was "totalitarian democracy" (Jacob Talmon, 1952). In his book, Talmon explores the reshaping of government into one in which social utility takes absolute precedence. While for Talmon the origins of totalitarian democracy are a merger of the interests of the individual and the state into a single *Volonté Générale*, in the Russian case, Rousseauist-Marxist and Jacobin Bolshevik lineages were only complementary to the recurring mode of the three successive liberation revolutions of self-integration of the country and its people. This system of governance was developing. During the Khruschevian period, the concept of an "All-Peoples State" stressing the democratic ideals of the nation of USSR was introduced. The Party was keen to maintain the integrity of the people and authority. A common slogan of the period was that the people and the Party are one.

Perestroika: More Democracy, More Socialism

In March 1985, Mikhail Gorbachev became the General Secretary of the Communist Party of the Soviet Union (CPSU). The key notions of *Perestroika* and *Glasnost'* entered public discourse at the time of the 27th Congress of the CPSU in February 1986. Slogans such as "more democracy, more socialism" and "back to Lenin" implied the embedding of democratic reforms in Soviet tradition. Yegor Ligachev, the party boss from Siberia whom Gorbachev had brought into the Politburo, asserted, "The Party expresses the profound essence of perestroika by the formula 'more socialism'" (*Pravda*, November 6, 1986).

The terms *democracy* and *democratization*, which had been used in various formulas by the Party in all periods, became essentially contested concepts. Typical of *perestroika* jargon was the "democratization" of everything. During the Congress, Gorbachev himself, as well as others, spoke about the "free revelation (*vyjavlenie*) of interests and the will of all classes and social groups" and about the "self-regulation and self-government of society." Though far-reaching in their implications, such statements gave no details of the institutional designs; a vague notion of "democratization" covered everything. When the legislative framework did appear, it was presented as a return to original Soviet institutions. Although the reforms marked a radical departure in the country's political structures, the constitutional arrangements adopted by the Supreme Soviet in November 1988 also left scope for traditional structures. Citizens would directly elect the People's Deputies, who would constitute the Congress of People's Deputies, which would in turn elect a bicameral Supreme Soviet. One third of the deputies would be elected, not by territory but by "public organizations" such as the Communist Party, the Young Communist League, and the trade unions. After a turbulent campaign, the first competitive elections were held on March 26, 1989, and more than 30 top party officials were defeated. The First Congress of the People's Deputies of the

USSR met in May 1989. The Extraordinary Fifth Congress met in September 1991 only to start the process of dismantling the USSR.

From a Crippled to a Sovereign Democracy

The end of the USSR in December 1991 marked a new stage in the fifth liberation revolution. It was the stage of consolidating new polities all claiming to become democracies. Various aspects of autocratic and democratic traditions as well as of their uneasy symbiosis manifested themselves in each of the cases of the 15 new independent states. In each case, the grand act of total emancipation turned into specific configurations of political institutions and practices. For the most part, those configurations were and continue to be highly ambiguous and contradictory. It was very symbolic that the key word for the Boris Yeltsin regime was *sovereignty*—not *samoderzhavie* this time, but still, integrity of power was very much the main political issue.

In the Russian Federation, interpretations of democracy and of new institutions of the country differed immensely, not only ideologically or partywise but also conceptually. There was a great gap between normative and pragmatic interpretations of democracy as such and the emerging democracy of the new Russia in particular. Often, if not typically, both visions were intricately intertwined. An apt example was the phase attributed to Deputy Prime Minister and actual head of the cabinet from 1991 to 1994, Yegor Gadar—"Russia is a crippled democracy, but it is a democracy at the end of the day."

Feelings that the newborn Russian democracy was inefficient and/or insufficient were very widespread in all political quarters of the country, but interpretations were quite different. That there was a great gap between the expectations and results achieved could not be denied. Some attributed the problems to the loss of an independent and consolidated power base of the regime. For others, the problem lay in the inability of the regime to stand up to normative ideals and best international practice. While the first view seemed more practical and consonant with Russian traditions, the second one was clearly impractical, wishful, and foreign. The first one gradually prevailed both in the minds of the authorities and the populace in general. The second view was characteristic of dogmatic "democrats" and zealous critics of the Yeltsin and, later, the Vladimir Putin regimes.

It was only natural that gradual consolidation of the political order and infrastructure of governance led to a reconceptualization of the constitutional framework of the Russian Federation as sovereign democracy (*suveryennaya demokratiya*). This formula was first introduced by Vladislav Surkov, the deputy head of the presidential administration, in his speech to the trainees of United Russia's Center for Party Personnel Training on February 7, 2006. Soon, the text was published as an article, "The Nationalization of the Future" in *Expert* magazine. It explained the notion as a proper synonym for "political competitiveness." Critics were quick to interpret the notion as a cover for exceptional democracy that would not tolerate any outside criticism. While this may be a secondary interpretation, the champions of sovereign democracy in fact insist that it highlights the great power status of Russia. Another possible, but unfortunately not widespread, interpretation is that of a sovereign (constitutional, etc.) state working to develop democracy.

It is very indicative that both Putin and Dmitry Medvedev deliberately refrained from public support of the concept. Rather, they voiced their reservations but agreed that the idea was worth pondering about. Still, it cannot be denied that with the notion of sovereign democracy, traditional priorities of political integrity have gained momentum. Although institutional and procedural aspects of accountability of authorities and of transparency remain grossly underdeveloped, sovereign democracy gives room for further promotion and anchoring of democratic institutions.

This short outline of Russia's leanings toward democracy shows that it is wrong to treat all its traditions as outright autocratic and democracy as something that had to start there from scratch. Russia's democratization story began as far back as the mid-19th century. Its heritage is rich albeit often tragic. Temptations to achieve radical breakthroughs to "pure" or complete integrity of authority and the people (far older than the democratization story itself) recurrently led to reestablishing the autocratic "integration" of Russia top-down. The present phase of sovereign democracy may be replaced by further moves toward more consequential accountability of authorities and popular political participation as long as the essential elements of the Russian democratic tradition are adequately used and pitfalls of radicalism avoided by the gradual growth of moderation and readiness to accept compromise solutions.

Mikhail Ilyin
MGIMO University
Moscow, Russian Federation

See also Democracy, Direct; Democracy, Types of; Democracy: Chinese Perspectives; Democratization

Further Readings

Benveniste, E. (1969). *Le Vocabulaire des institutions indo-européennes* [The vocabulary of Indo-European institutions] (2 vols.). Paris: Minuit.

Kahn, J. (2002). *Federalism, democratization and the rule of law in Russia*. Oxford, UK: Oxford University Press.

Kavelin, K. D. (1989). *Our mental constitution.* Moscow: M. Pravda.

Lincoln, W. B. (1981). The problem of Glasnost' in mid-nineteenth century Russian politics. *European Studies Review, 11*(2), 171–188.

McDaniel, T. (1996). *The agony of the Russian idea.* Princeton, NJ: Princeton University Press.

Petro, N. (1995). *The rebirth of Russian democracy: An interpretation of political culture*. Cambridge, MA: Harvard University Press.

Pivovarov, Y., & Fursov, A. (2001). The "Russian System" as an attempt to interpret Russian history. *Polis, 4*, 37–48.

Riasanovsky, N. V. (2005). *Russian identities: A historical survey*. New York: Oxford University Press.

Talmon, J. L. (1952). *The origins of totalitarian democracy.* London: Secker & Warburg.

von Haxthausen, A. (1847–1852). *Studien über die inneren Zustände, das Volksleben und insbesondere die ländlichen Einrichtungen Russlands* (3 vols.) [Studies about the inner affairs, the life of the people and, in particular, the rural institutions of Russia]. Hannover (Vols. 1 & 2)/Berlin (Vol. 3), Germany: Hahn.

Democratic Consolidation

Democratic consolidation may be described as the process of defining and firmly establishing the essential characteristics (and adjusting the secondary ones) of the structures and norms inherent in a democratic regime, which comes about also, but not exclusively, with the passage of time. Institutions, procedures, practices, customs, and routines are defined and adapted, and at the same time, the structures and regulations for the peaceful resolution of conflicts become accepted, thus strengthening the regime's legitimization. In short, consolidation is one of the possible outcomes arising out of the introduction of democracy: This may result not only in a fully functioning democracy but also in its bare survival or in a new crisis of the political system and a subsequent shift toward authoritarianism. This entry defines democracy and clarifies the principal meanings of consolidation used in this analysis. After a brief discussion of some criticisms of the concept, the entry examines the empirical dimension and the mechanisms through which democracies can be consolidated. Finally, following an analysis of the factors that favor this process, some concluding considerations are offered.

Definitions

Defining the concept requires first of all (a) identifying when a government can claim to be at least minimally democratic and (b) outlining the different types of democracy along a continuum running from partial forms of democracy to more complete ones. As far as the first objective is concerned, one can start with the well-accepted notion of polyarchy, whereby a political regime is democratic when key civil and political rights are protected and elections are free, transparent, and competitive. The minimum forms of democracy have been defined as "electoral democracies": In these democracies, some of the procedural requisites of polyarchy are respected, but elections are only partially free, and the chances of an opposition victory are limited. For a regime of this type to claim to be democratic, however, election results must reflect the will of the voters, and in particular, it must be possible to remove unpopular politicians from power. This necessitates an open electoral arena, with substantial freedom for parties and candidates to publicize their ideas and canvass votes, along with guarantees that they may speak, publish, meet, organize, and move freely within the country to pursue their objectives peacefully.

In some historical and geographical contexts, democracies have achieved a certain level of maturity, the government and public administration operate better, the judicial system is more independent and effective, parties and interest groups are better structured, the democratic political culture is more robust, and civil society is autonomous and vibrant. Likewise, governments react responsibly to requests from voters and social organizations and endeavor to represent the interests of all—including marginal and weaker groups and ethnic, cultural, and religious minorities. In this case, one can legitimately speak of "liberal democracies."

To consolidate a democracy implies strengthening it. But democracy can be strengthened in different ways. Originally, it was thought that strengthening a democracy was synonymous with making it "secure," extending its "life expectancy" and safeguarding it from the danger of authoritarianism. Subsequently, this interpretation was widened to incorporate a series of other "missions." The resulting long list thus includes neutralizing antisystem actors, building a party system, organizing economic and social interests, stabilizing electoral laws, decentralizing state power, reforming the judicial system, and solving problems of poverty and underdevelopment. The new objective, therefore, seems to be that of getting as close as possible to the organizational forms and political and social performances of advanced Western democracies. In this sense, strengthening a democracy means deepening its characteristics and realizing its full potential.

Seen in this light, democratic consolidation takes on two very different meanings. The first is substantially negative: It implies avoiding authoritarianism—that is, moving from an "electoral democracy" to a new authoritarian regime—or preventing the erosion of democracy and its recession from advanced forms to other more precarious ones. In the first case, the danger is that of a sudden death, that is, of democratic breakdown; in the second, the risk is of impoverishment and a slow death, of a gradual unraveling of democracy and an almost imperceptible return to authoritarian political forms. The second meaning of consolidation is positive. Consolidating democracy in this sense means bringing it to fruition, progressing from an initial form to other more complete forms, from "electoral" forms of democracy to a democracy in which civil and political rights are more numerous, more detailed, and better protected. According to many, social and economic rights should also be fully protected, and consolidating democracy thus means creating an "advanced democracy," which represents the ideal goal of more progressive political practices.

Which of these interpretations is preferable? It seems appropriate to limit the term to the first set of meanings, the negative ones and in particular to that of survival: Consolidation should imply only the expectation of the regime's endurance. Some find the other definitions problematic, since democracy and its identity cannot be statically and definitively fixed, and its contents are in perennial movement, always open to new clarifications. Hence, it would be analytically impossible and futile to attempt to define once and for all what constitutes a "good democracy" and a "finally consolidated" democracy. Given the confusion surrounding it, the concept of positive consolidation associated with this interpretation is equally empty. Those scholars who believe that one can study democratic performance and its potential for full realization through other concepts, such as the quality of democracy, have reached similar conclusions: For them, too, the typical sphere of consolidation remains the original one, rooted in the prospects of democracy's maintenance and mere survival rather than the possibility of full democratic realization, which would again involve the concept of quality.

Although they are always changeable and undefined, the specific contents of democracy are not entirely indeterminate and are relevant to the prospects for consolidation. In fact, it is no coincidence that the consolidation of modern democracies requires the crafting of mutually acceptable rules that concern crucial themes such as negotiation between capitalists and workers, guarantees of different forms of property, and assurances concerning levels of social protection to be established or defended. In more general terms, democracy enjoys better consolidation prospects when the major political actors have a stake in its survival. In this setup, actors outside the electoral arena, especially the military and entrepreneurs, receive guarantees that their vital interests will be respected. Likewise, political parties and movements not in government must have realistic chances of taking power through elections. Only when an agreement of this type exists can they too, over time, become accustomed to democracy. Once they realize that their initial fears were mainly groundless, they can transform lukewarm acceptance of the new regime into full support.

Criticism

Some political scientists have (at times sharply) criticized the concept of democratic consolidation and deemed it unacceptable on account of its strong teleological flavor. Cases that have not achieved full institutionalization are seen as enduringly unconsolidated. Such a view presupposes that there are factors working in favor of increased consolidation or institutionalization but that countervailing obstacles obstruct a process of change that otherwise would operate unfettered. The current existence of many countries that have been in a state of failed or partial consolidation for decades, however, casts the existence of this linear progression into doubt. To claim that consolidation occurs only when the behavior of the main political actors is adapted to the formal rules of democracy would be to ignore the experience of many countries, such as India or Italy, that are widely considered to be consolidated democracies but in which particularism and patronage remain widespread. Others, finally, find no trace of a specific empirical dimension of the concept, which should be observable if the chances of a democratic regime's survival over time were increased simply because it has already survived for a sufficiently long period.

Empirical Indications of Consolidation

Many political scientists have focused their attention on this dimension, wondering, at first, how a consolidated democratic regime can be empirically recognized. Even if this analysis is limited to the prospects of regime survival, as discussed above, the answer to this question is problematic. First and foremost, it is important to note that some have studied the behavior of political actors, and others their attitudes. The former in particular have concentrated on antidemocratic behavior as a symptom of the weakness of a democracy and have tried to determine how and when such behavior violates its basic rules. Principally, democracy is challenged when political actors try to achieve their goals through violence. These acts include a series of actions ranging from the systematic destruction of public or private property to the assassination of political adversaries, and to attempts to forcibly overturn the government and intimidate voters and candidates. The refusal to recognize the validity of genuinely free political elections can also trigger an irreversible democratic crisis, for example, by deciding not to participate in voting, stopping others from participating, or manipulating electoral results at will. Finally, consolidation implies that political actors must respect the laws and the constitution, and accept the agreed rules of political behavior. When the government violates the established rules while legislating, enforcing laws, interpreting them, or resolving conflicts, and especially when these violations become customary, the survival of democracy is severely at risk.

Having clarified these essential points, many other questions nevertheless remain unanswered. For example, which levels of antidemocratic behavior are sufficient to trigger a democratic crisis and can therefore be considered a symptom of de-consolidation? It is obvious that the higher the levels, the greater the risk for the new democracy, but it is also possible that relatively low levels can accumulate over time to produce a crisis that may be unexpected. And it is perfectly possible that a democratic crisis will serve to bolster the new democratic regime rather than destroy it. This was the case, for example, with Colonel Antonio Tejero Molina's attempted coup d'état in Spain in 1981. Finally, consolidation can be considered authentic only if it guarantees democratic survival when adverse external conditions are present. It is essential, for example, that a governing party peacefully accepts giving up power in the event of an opposition election victory. The turnover-in-government test reflects this rationale, and if turnover takes place twice, then the test is even more stringent: In this latter case, not only are we then certain that two large democratic parties have been formed but also that elite groups and public opinion fully accept the functioning of the democratic mechanism, as they have proved that they are willing to change the government, but not the democratic regime, when things have gone badly.

A second empirical indicator of consolidation concerns the attitudes of political actors as regards both normative and cognitive aspects. Democracy is consolidated normatively when its citizens consistently give it genuine support and value it as an intrinsic good, not merely an instrumental one. For many, this transformation signals the conclusion of the process: Although democracies can survive at relatively low levels of consensus, it is undoubted that marked legitimacy provides significant help in tackling and overcoming moments of crisis, including acute ones. The most widely used tool to determine whether, and to what extent, the citizens of new postauthoritarian regimes appreciate democracy is opinion polling. Undoubtedly, this is still a fragile tool, especially if polls consist of generic questions that are difficult to interpret with any certainty, such as those concerning public preference for democracy over other political regimes. This is not only because alternative experiences to democracy in some cases are only a distant memory but also because the dimensions and experiences linked to democracy are too numerous, and in part contradictory, to allow for a single reply. Among the attitudes to consider are the democratic expectations of political actors, that is, the extent to which they think democracy will survive—a cognitive aspect relating to personal evaluations. Again, as in the case of public consensus, opinion polls are a significant, though not always complete, tool to analyze some of their empirical features.

Mechanisms of Consolidation

Having attempted to define the concept in its theoretical and empirical components, this entry now looks at how it operates. Which processes accompany the consolidation of democracy? Who are the actors and what strategies do they employ? Consolidation implies the definition of relations between representative and governing institutions and civil society. These relationships develop in two directions: from the bottom upwards, that is, from society to the institutions, and from the top downward, in other words, from the newly created institutions through intermediate structures to society. In the first case, it is civil society that backs the regime through a process of legitimization and support. It is worth noting that some scholars claim that the crucial aspect for the survival of a regime is not active support but rather the lack of organized alternatives. Consensus, therefore, may often be only acquiescence and passive acceptance of the status quo. However, no matter how one defines the foundations of legitimization, it must be stressed that, if legitimization is widespread, if all political organizations are integrated and involved in accepting and supporting democratic institutions, then an important step toward democratic consolidation has already been taken in practice. By contrast, where support is limited, consolidation is only possible through the opposite mechanism, known as "anchoring," which operates from the top down.

The key "anchors" of democratic consolidation are political parties and their organizations (parties may be more or less institutionalized), their gate-keeping function (with which they control the access of interest groups to the decision-making arena and establish priorities among different demands), their links of patronage (particularistic relationships with large and small private entrepreneurs, intellectuals, and ordinary citizens), and neocorporatist agreements. In sum, political parties exercise their function of political "anchors" by organizing, integrating, and sometimes even controlling general civil society and the specific sectors that make it up, consequently helping strengthen the democratic political regime. The crucial point is that the process of consolidation and the particular mix of its two components (legitimization and anchorage) give rise to different consolidation formulas in different countries, thus enabling one to determine the type of democracy that derives from it. For example, in Italy, legitimization of the democratic regime created after World War II was limited, and the parties established a position of domination over interest groups, giving rise to consolidation of the party-based type. By contrast, in Spain after Franco's death, the parties were generally weaker than in Italy and were more neutral toward interest groups. However, in the period after democratization, the new regime acquired strong legitimacy, partially due to the role played by the party elites: Spain is therefore an example of consolidation through the elite. It should be noted, though, that especially if they have contributed to democratic consolidation, the principal "anchors" may not be very adaptable to the changing circumstances of political life, and a series of factors—long permanence in power, lack of turnover, absence of coalition partners in government—may throw the new democratic regime, and the prevalent type of consolidation, into crisis. This in turn may lead to instability and produce the contrary phenomenon of "disanchoring" of civil society from politics.

Factors Influencing the Success of Consolidation

The conditions that favor the success of democratic consolidation remain to be discussed. The list is long; here a general outline is provided, distinguishing between political, economic, and international scenario-related factors. Among the first, the nature of the previous authoritarian regime is significant: Some scholars argue that if political repression was intense or if the authoritarian regime collapsed after military defeat, the new democratic government will enjoy "inverse" legitimization thanks simply to having replaced a widely discredited government; however, this effect does not last long. Others are of the view that democratic structures are more easily reinforced if less violence has been used and the attitude of the defeated authoritarian government is one of compromise and negotiation with the democratic forces.

Previous democratic experiences may be of more significance: If some democratic institutions were maintained during the authoritarian interval, consolidation is more likely to be successful, compared with a scenario in which such institutions have to be created *ex novo*. In the first case, the prior democratic experience greatly facilitates the restoration of previous alliances between parties and groups and the reemergence of associations, trade unions, and other organizations that support democracy. Indeed, during the 20th century, few countries succeeded in introducing a stable and lasting democracy at the first attempt. Nonetheless, more recent experiences, involving the progressive democratization of countries with little or no previous democratic history, may open new and interesting scenarios for the survival of democracy in such cases.

Certain structural elements, such as some degree of ethnic and religious homogeneity and the absence of deep ideological divisions or severe economic and social inequality, appear to favor the emergence of a strong political culture, which is another important determinant for consolidation. These characteristics are also associated with a robust civil society. Democratic consolidation is thus held to benefit from associations and institutions that are strongly rooted in civil society, provided that these try to achieve their own interests also in opposition to political power, in turn restricting the desire of the latter to expand disproportionately. In order to be effective, these associations must be strong, centralized, and willing to mobilize for political objectives; otherwise, the government will find it easy to deal with those opposing it, whether by intimidating them or buying their services.

Among the political institutions that facilitate the process, parties are undoubtedly of central importance; this entry has already mentioned their role in the development of consolidation, as agents promoting legitimacy bottom-up and, if this fails, anchoring it top-down. Parliamentary systems are thought by some scholars to be more favorable to democratic consolidation than presidential ones, since the former reduce the harsher aspects of political confrontation through the creation of party coalitions necessary to form a government and because they provide a better balance between the powers of the president and those of the head of government. Electoral systems are also the subject of debate, with the question posed of which more effectively strengthens democracy: a proportional system, which better represents the claims of economic, social, local, and ideological minorities, or a majoritarian system, seen as more capable of delivering a clearer and more credible governing alternative? Although the empirical evidence remains somewhat ambiguous, the prevailing current of thought is that parliamentary systems are better in this sense than presidential ones, especially if the latter are combined with a proportional party system, which multiplies its limitations.

A second group of conditions that facilitate democratic consolidation are economic in nature and involve two possible scenarios. Reaching certain levels of development is thought to reflect the fact that a democratic regime has been consolidated and that it no longer runs a significant risk of lapsing into authoritarianism. In Argentina, the per capita income in 1975 was above $6,000, still one of the highest levels of wealth ever associated with a democracy that was overturned. However, wealth is not necessary for consolidation: Even poorer countries can maintain a democratic regime if they can generate strong economic growth, accompanied by moderate and stable levels of inflation and if they extend the new wealth to the majority of the population. The probabilities of a democracy dying, in particular, also appear to vary inversely with economic inequality, in line with what was already suggested in the 1960s. In brief, it is reasonable to hypothesize that a modern industrialized economy, in a complex relatively egalitarian society with an educated population, favors the maintenance of democracy. One of the implications of this analysis is that while economic crises destabilize all democracies, they hit poor democracies particularly hard since they are more vulnerable to weak economic performance. Consolidation is also helped by the presence of an adequate social welfare system. If democratic governments succeed in reducing social and economic inequality through appropriate social policies, disorder and rebellion in the poorer countries will be reduced, and in all cases, the sensation of the legitimacy and efficacy of the new regime will be strengthened. Indeed, recent studies have documented the growth in social spending during the phases following the introduction of democracy in several of the third wave of democratic transitions, with taxation often having been increased to finance and extend welfare systems, reduce poverty levels, improve medical assistance, and expand education.

The international scenario and external actors may also be of considerable importance. Think, for example, of the role played by Federal Germany in the democratization and political consolidation of East Germany and that of the European Community (later the European Union) with regard to Southern and Eastern European countries. Especially in the cases of more recent democratizations, the international system has either facilitated or hindered existing democratic tendencies: in the former cases, acting to support them when the democratic context was widespread and solid, and in the latter cases, acting as an obstacle when prevalent governments and political regimes were hostile or indifferent to democracy. It is worth noting that some paradoxical and unexpected effects, however, for example, in some cases, the excessive encouragement of democratization by powerful international actors—before adequate economic and social conditions have taken root in the country in question—have exacerbated problems of consolidation. There is also a danger that regimes to which aid is destined are satisfied with acquiring a veneer of international legitimacy, without actually modifying the authoritarian nature of domestic power management.

Conclusion

The concept of democratic consolidation is complex, problematic, and to some extent contradictory. This makes it all the more important to clarify and circumscribe the meaning of the term. It has been suggested by some scholars that the concept be limited to its original significance, that is, the survival of democracy to avoid the risk of an authoritarian regression. As proposed above, this implies limiting its application to the "negative" sphere: preventing democratic collapse or protecting democracy from a slower death through erosion. First and foremost, this choice enables us to reestablish the many meanings associated with the concept, which can then be used more consciously and consistently. The other sense in which it is widely used—that of completing and satisfying the many "unfulfilled promises" of democracy, to use Norberto Bobbio's well-known expression—calls for an analysis of the goals and criteria held to be indispensable for achieving "full" and desirable democracy. This type of analysis appears to belong to studies on the "quality" or "full achievement" of democracy and should be developed in that context. In the wake of the more recent experiences of democratization, studies of this type are now increasing sharply. They are still only at the beginning, and their preliminary results require verification; however, they represent a coherent means of completing the analysis of the functioning and performance of democracy that originated in studies of democratic consolidation.

Davide Grassi
Università degli studi di Torino
Turin, Italy

See also Democracy, Quality; Democracy, Theories of; Democratization; Legitimacy; Transition

Further Readings

Bobbio, N. (1984). *Il futuro della democrazia* [The future of democracy]. Torino, Italy: Einaudi.

Huntington, S. P. (1991). *The third wave: Democratization in the late twentieth century.* Norman: University of Oklahoma Press.

Linz, J., & Stepan, A. (1996). *Problems of democratic transition and consolidation*. Baltimore: Johns Hopkins University Press.

Morlino, L. (1998). *Democracies between consolidation and crisis: Parties, groups and citizens in Southern Europe.* Oxford, UK: Oxford University Press.

O'Donnell, G. (1996). Illusions about consolidation. *Journal of Democracy, 2,* 34–51.

Przeworski, A., Alvarez, M. E., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development.* Cambridge, UK: Cambridge University Press.

Rustow, D. (1990). Democracy: A global revolution? *Foreign Affairs, 4,* 75–91.

Schedler, A. (1998). What is democratic consolidation? *Journal of Democracy, 2,* 91–107.

Schedler, A. (2001). Measuring democratic consolidation. *Studies in Comparative International Development, 1,* 66–92.

Schmitter, P. (1998). Some basic assumptions about the consolidation of democracy. In T. Inoguchi, E. Newman, & J. Keane (Eds.), *The changing nature of democracy* (pp. 23–36). New York: United Nations University Press.

Democratic Peace

Democratic peace refers most specifically to the proposition that democratic states have not fought and are not likely to fight interstate wars against each other. It refers more generally to the notion that democracy has an important pacifying impact on international politics, whether by making individual states less warlike, by creating peaceful relationships for pairs of states, or by making the entire globe more peaceful. It is a proposition that lies at the heart of the American academic field of international politics, born in the wake of a world war ostensibly fought to make the world safe for democracy. Though its philosophical roots go back to Immanuel Kant and Thomas Paine, in its contemporary form, democratic peace focuses most intently on democracy's impact on interactions within pairs of states. This entry reviews research on that impact, its theoretical bases, major criticisms of it, and the most important responses to those criticisms.

Impact of Democracy

Recent interest evoked by democratic peace is the result, in an important part, of the strikingly simple claim that no two democratic states have ever fought a war against each other. This is not a trivial claim. Though interstate wars are far too common, statistically speaking, they are rare events. In most years, 99% of the pairs of states in the international system avoid fighting wars against each other. For the rate of warfare among democratic states to be *significantly* different from that for states, in general, the number of wars between democratic states must be at least close to zero.

So critics of democratic peace point out exceptions to the alleged rule about democratic states having universally peaceful relationships with each other. The War of 1812 between the United States and Great Britain, the U.S. Civil War, the Spanish American War in 1898, even World War I, the official state of war between Great Britain and Finland in World War II, the war between Lebanon and Israel in 1948, and the military conflict between India and Pakistan in 1999 are among the most frequently mentioned wars on this list.

Resolving the debate about whether any or all of these are actually exceptions to the democratic peace rule obviously must involve definitions of "war" and "democracy." A specific definition of war widely adopted by researchers focusing on quantitative analysis of evidence regarding its causes specifies that an interstate war involves military conflict between independent states leading to the deaths of at least 1,000 soldiers. This definition has been adopted by most analysts conducting systematic empirical evaluations of hypotheses regarding the causes of war for several decades.

Definitions and measures of democracy are probably more contestable. Most of the research on democratic peace has adopted a numerical threshold based on data consisting of annual scores rating states on a continuum from fully autocratic to entirely democratic. These thresholds are inescapably arbitrary to some extent. In fact, all definitions contain an important arbitrary element. Furthermore, to evaluate the validity of the statement that democratic states never fight wars against each other, all states must be sorted into "democratic" or "not democratic" categories. Obviously, states do not naturally or clearly fall into neatly exclusive or exhaustive categories of that kind. So, ultimately, the proposition that needs to be evaluated is "states that are sufficiently democratic never (or rarely) fight wars against each other." How much democracy is enough, and how can that level of democracy be identified, are crucial questions with no answers that will generate universal consensus. One characteristic receiving some attention focuses on the ability of a state to stage elections in which executive leadership of the state passes from one independent political party to another, different independent political party.

Theoretical Frameworks

The debate about which attributes commonly thought of as "democracy" possess the most effective pacifying impact leads inevitably to a more general issue regarding the theoretical basis for the democratic peace proposition. This debate shifts the focus from questions about whether it is true that democratic states have fought wars or become involved in militarized disputes at a lower rate than other kinds of states to a rather different issue: Assuming an acceptance of the claim that jointly democratic states do have a historical record of conflict that is distinctive in the way the democratic peace theory would suggest, what has brought about this record? One prominent answer focuses on the potential cultural or psychological impact of democracy on states and their leaders. Political leaders in democratic states are accustomed to resolving political debates and disputes in a peaceful, rule-based fashion. Leaders of democratic states can be expected to bring attitudes and expectations about conflict resolution cultivated in domestic political processes to interactions with their counterparts in other states. For their own protection, democratic leaders bring quite different attitudes and expectations to their interactions with leaders of autocratic states. They find such leaders ethically and morally suspect from the outset. And they have no expectation that conflicts with them can be resolved in a peaceful manner. But if the leaders of other states are democratic, leaders of democratic states anticipate that disagreements can be resolved in a peaceful manner.

The second main theoretical argument about the potential pacifying effect of democracy focuses on the structure of domestic political systems. An important starting point for such arguments posits that political leaders of all states place the highest priority on staying in power. Analysts point out that to initiate military interventions or attacks, leaders of democratic states need to persuade various elements in the government, interest groups, and even the general population to support such ventures. This process, relatively cumbersome compared with that faced by autocratic leaders, makes democratic states more hesitant and cautious about getting involved in military conflict. Then, too, democratic leaders are more vulnerable to being removed from office if they become involved in a war, or even a less serious militarized conflict, and they lose that war or conflict. Since democratic states tend to win the wars in which they become involved, democratic states are, everything else being equal, opponents to be avoided in military conflicts. Autocratic leaders tend to be not so cautious and not so anxious to win the wars in which they are entangled. All they need to do to maintain themselves in power is to satisfy or appease their relatively small winning coalitions. And since those winning coalitions are small, supporters of dictators tend to be loyal. They understand that they might well be excluded from the small winning coalition supporting any successor regime.

Democratic leaders must satisfy a much larger coalition. Therefore, they depend more crucially on providing public goods to that larger coalition, such as victories in war. If a war is lost, they probably cannot provide enough private goods to a sufficiently large portion of their coalition to maintain themselves in power. This is especially the case because supporters of democratic leaders tend to be much less loyal than supporters of autocratic leaders. The winning coalition of any successor in a democratic regime must necessarily be large, at least approaching one half of the population involved in selecting leaders—that is, the entire electorate. In other words, in political systems based on relatively large winning coalitions, those who desert the coalition supporting the current leader also have a reasonable chance of joining the winning coalition of the successor. All of these considerations make democratic states cautious about getting involved in wars or militarized conflicts, especially against other democratic states.

Criticism and Response

Some critics of both these approaches to explaining the alleged lower rate of warfare and militarized conflict among democratic states acknowledge that the statistical pattern exists. They argue, however, that the correlation between democracy and peace for pairs of states is not the reflection of a causal linkage. One argument has it that it is actually peace that produces democracy. That is, states that for whatever reason come into existence and/or exist for prolonged periods of time in an environment of peace are more likely to be democratic. But this does not prove that democracy causes peace.

One of the most prominent antidemocratic peace arguments points out that most of the historical examples of democratic pairs of states avoiding wars against each other come from the decades following the onset of the Cold War. Democratic states in that era, it is argued, avoided wars against each other not because they were democratic but because they shared the unifying threat of a powerful common enemy in the Soviet Union or the communist coalition of the Soviet Union, the People's Republic of China, and various allies.

However, several statistical analyses of the relationship between democracy and peace control for the impact of alliance ties and still find that democracy apparently leads to peace. And, it is not true that the potentially unifying impact of opposition from the communist world was universally effective. That opposition did not prevent, for example, a war between El Salvador and Honduras in 1969, military conflict between Greece and Turkey over Cyprus in 1974, or the clash between Great Britain and Argentina over the Malvinas/Falkland Islands in 1982. In each of these cases, at least one of the antagonists was not democratic. Furthermore, the communist world was confronted by a capitalist/democratic world even more imposing than itself in terms of military and industrial capabilities, and the allegedly unifying impact of that opposition did not prevent the Soviet Union from invading Hungary, Czechoslovakia, or Afghanistan, the serious clashes between China and Vietnam, and Vietnam's attack on and occupation of Cambodia.

Democratic peace ideas influenced the foreign policies of the Clinton and Bush administrations during the 2 decades after the end of the Cold War. Especially in the latter case, some foreign policy initiatives ostensibly based on those ideas had problematic outcomes. One important criticism of the democratic peace literature is that it has never laid out clearly the policy implications of its findings, even though it is also true that few if any democratic peace advocates, for example, endorsed the U.S. intervention in Iraq in 2003. In short, democracy may have an important pacifying impact on relationships between states. How policymakers might best take advantage of that impact is not always clear.

James Lee Ray
Vanderbilt University
Nashville, Tennessee, United States

See also Conflict Resolution; Cooperation; Dictatorship; International Relations, Theory; Kant, Immanuel; Liberalism in International Relations; Peace; War and Peace

Further Readings

Gowa, J. (1999). *Ballots and bullets: The elusive democratic peace*. Princeton, NJ: Princeton University Press.

Ray, J. L. (1998). *Democracy and international conflict: An evaluation of the democratic peace proposition*. Columbia: University of South Carolina Press.

Rummel, R. J. (1979). *Understanding conflict and war: War, power, peace* (Vol. 4). Beverly Hills, CA: Sage.

Russett, B., & Oneal, J. (2001). *Triangulating peace: Democracy, interdependence, and international organizations*. New York: W. W. Norton.

Weart, S. (1998). *Never at war: Why democracies will not fight one another*. New Haven, CT: Yale University Press.

Democratization

Democratization can be defined as two different political processes:

1. Democratization is the transformation from a nondemocratic regime to a democratic political regime. It is a transition from nondemocratic to democratic types of political regime and involves regime change.
2. Democratization is the process of political transformation from an electoral or partial democracy toward a full or consolidated democracy. It is a transition between different degrees of democracy within a specific democratic political system.

This entry covers both meanings, by discussing the actors who may bring democracy, the factors of and impediments to democracy, and historical and future perspectives of democracy and the democratic trend.

Agency Factors of Democratization

The main theory with emphasis on "agency," as developed in sociological theory, is the transition approach, which was proposed by Guillermo O'Donnell and Philippe C. Schmitter (1986) and further developed by Gretchen Casper and Michelle Taylor (1996). This agency approach analyzes the process of transition toward democracy with political elites as the main actors and as the crucial and critical factor of democratization. One of the conditions that help initiate a transition to democracy in an authoritarian regime is if the ruling elite splits into factions with opposing interests. This is likelier to happen in more developed societies whose complexity creates multifaceted regime coalitions that are not as easily held together. Rifts within the ruling elite are also more likely when there is a mounting legitimacy crisis due to economic setbacks, unfulfilled policy promises, and failures in crisis management.

In heterogeneous regime coalitions, legitimacy crises encourage elite splits because they create an opportunity for some elite groups to try to strengthen their position in the regime coalition by pursuing a reform strategy that they hope will bring them popular support, thus regaining legitimacy. Accordingly, many transitions to democracy have been instigated by the emergence of a reform camp within the regime. Typically, the reformers initiate a liberalization program that opens a space for criticism and alternative voices. As a result, opposition groups surface from the underground and, in many cases, advance further claims for democratization. If the opposition groups remain moderate in their methods (avoiding violence) and demonstrate their readiness for compromise but at the same time muster widespread public support, a negotiated transition to democracy becomes possible.

The emergence of opposition to a regime does not always result from an elite-initiated opening process. Sometimes, policy failures lead to spontaneous manifestations of widespread mass opposition, launching a legitimacy crisis that impels an intra-elite reform camp to surface and engage in negotiations with the opposition. Again, this configuration of events often leads to negotiated, or "pacted," transitions.

The institutional basis of a given authoritarian regime is an important factor in this context because different types of authoritarian regimes show different vulnerabilities to democratizing pressures. For instance, the weakness of military regimes is that they lack an ideological mission that legitimates them on a long-term basis. Usually, they take power as crisis managers, so their justification is—often explicitly—only temporary. The legitimacy of military regimes is relatively easily questioned, either because the junta fails to manage the crisis, in which case its justification lacks credibility, or because things run smoothly, in which case the need for crisis management becomes obsolete. One obvious advantage of military regimes is that they control the means of coercion, so they can silence emerging opposition by brute force. But confronted with widespread mass opposition that proves resilient even in the face of oppression, the loyalty of the troops may erode if they are ordered to turn on peaceful protestors. On the other hand, even though military regimes sometimes exit quickly from power, they also easily return, as the repeated oscillations between military and civilian rule in countries such as Turkey, Pakistan, or Thailand demonstrate.

Personalistic regimes put all their eggs into the basket of the central ruler's charisma. Accordingly, when the ruler dies, there is an opportunity for political change, as has been clearly demonstrated in Spain. Whether or not this opportunity is used for a transition to democracy then depends on the power balance between prodemocratic and antidemocratic forces and their relative support among the population.

One-party regimes, whether leftist or rightist, profit from a more strongly institutionalized power basis. These regimes usually have an ideological mission that inspires their existence and provides legitimacy. It generally takes longer and is a bigger challenge for a potential opposition group to erode the ideological basis of one-party regimes. One strategy that proved successful in the former communist bloc is to demonstrate that a regime betrays its very own ideals. When communist countries signed the human rights declaration in the Final Act of the Conference on Security and Co-operation in Europe (CSCE) while refusing to respect these rights in practice, civil rights movements such as the Charta 77 effectively publicized this contradiction and in doing so helped erode the regime's legitimacy. Eventually, the legitimacy crisis went so deep that even within the communist parties no one believed any longer in the regime's ideals. In this situation, reform camps surfaced in a number of communist parties (most notably in the Soviet Union and Hungary), together with organizations outside the party that were opposed to the regime, once Mikhail Gorbachev's nullification of the Brezhnev Doctrine in 1998 made this seem like a viable strategy in Central and Eastern Europe.

Splits in the ruling elite are important because they give leverage to domestic as well as international actors, enhancing their bargaining options to push a democratization agenda through. The leverage that international actors have in pushing for democracy increases insofar as a country depends on international aid. In some cases, dependence on international assistance can be so strong that external powers can trigger democratization even in the absence of a prodemocratic opposition group within the country. In the extreme case, democratic powers can enforce democratic institutions by military intervention as was attempted in Afghanistan and Iraq. But externally triggered processes of democratization are unlikely to penetrate very deep unless there are strong domestic forces inside a country.

When parties of the ruling elite are unified to sustain an authoritarian system, the transition to democracy is less easily achieved, particularly if the regime is able to isolate itself from international democratizing pressures. In such cases, the chances to democratize depend strongly on whether a prodemocratic opposition group emerges, how massive it grows, and how skillfully it uses its repertoire of elite-challenging actions. If the opposition can mobilize support from all layers of the population, if it is able to demonstrate this support, and if it remains resilient even in the face of oppression, loyalty to the regime erodes, even among the armed forces, undermining the regime's repressive capacities. Thus, massive, determined, and well-organized opposition groups can overcome the resistance of the ruling elite to democratization. If, however, the opposition remains limited to isolated sectors of society, is unable to demonstrate popular support across the country, and cannot remain resilient in the face of repression, its chances of success will be limited.

To a considerable extent, then, democratization is a matter of the skills and virtues of mass opposition leaders. It matters how willing and able they are to advance claims that resonate with many people, to mobilize resources for popular campaigns, and to make use of the full set of ruling elite–challenging actions even in the face of repression. Tactical and strategic factors, such as the presence of skilful political dissidents, benevolent reforms by the ruling elite, and international assistance, are important, but when it comes to full democratization, these factors can hardly compensate for deficiencies in the development of ordinary people's capabilities and motivation to practice democracy.

Factors of Democratization: Economic, Social, and Cultural Modernization

Mounting and sustaining a prodemocratic opposition against authoritarian rulers require that societies embark on a process of human empowerment that develops organizational skills and resources among the population. It also requires that substantial segments of the population give high priority to democratic freedoms, making them able and willing to struggle for democratic institutions. Ordinary people's readiness to struggle for democratic freedoms is necessary for deep democratization to be attained, for authoritarian leaders are unlikely to surrender their powers unless they are pressured to do so.

Paramount is a type of economic development that is knowledge driven and distributes human social and political resources widely throughout society rather than concentrating them in small minorities of the population. Industrialization and the rise of the knowledge society equip growing segments of the population with the material means, intellectual skills, and social opportunities needed to mount effective pressures on elites. As a consequence, ordinary people's action repertoires expand in ways that make the value of democratic freedoms intuitively obvious, giving rise to emancipative worldviews that value freedoms highly. These long-term developmental factors enhance a society's ability and willingness to struggle for democracy.

The crucial and critical importance of economic, social, and cultural modernization has been at the core of a modernization theory of democratization. The main author of such a theory of modernization as the most important factor of democratization is Seymour Martin Lipset (1994). This modernization theory of democratization was expanded and refined by Larry Diamond (1992). Hence, this theory of democratization can be called the Lipset-Diamond theory of democratization.

The human empowerment approach of democratization is a hybrid structure–agency theory of democratization and has been developed by Ronald Inglehart and Christian Welzel (2005).

External Threats and Group Hostilities as Impediments to Democracy

Various situations can arise that prevent developmental factors from giving rise to the prodemocratic effects associated with them. Perceptions of external threats and internal group hostilities tend to be detrimental to democracy because they diminish tolerance of opposition—a basic principle of democratic organization. External threats help leaders use "rally around the flag" strategies that silence inner opposition. Group hostilities do the same within groups, closing ranks around leaders and silencing opposing views.

Involvement of a country in an enduring international conflict can undermine democratic institutions because they provide a sense of being threatened that allows skilful leaders to present suppression of the opposition as crucial to the nation's survival. Similarly, internal group divisions are not inherently threatening to democracy, but ethnic, linguistic, religious, and other easily discernible group divisions can be manipulated to foment support for authoritarian leaders. Extremist leaders almost always mobilize support by playing on group hostilities. Thus, democracy has historically been more easily established and consolidated in societies that are relatively homogeneous culturally and relatively egalitarian economically.

Regardless of whether such hindering factors are present, deep democratization requires that a society's people acquire the capability and motivation to struggle for the freedoms that define democracy. This is because democracy is a socially embedded phenomenon, not just an institutional machine that operates in a vacuum. Shallow democratization involves crafting institutions, but deep democratization involves the development of supportive values and skills among large segments of a society.

Democratization and Transition in a Long-Term Historical Perspective

How can we understand the fact that democratization processes in separate countries cluster into coherent and sweeping international waves, behaving as if they were centrally coordinated by a master agent when in fact neither that master agent nor central coordination of the international waves exists? The answer is that evolutionary forces are at work that go beyond the awareness and control of even the most powerful elites. These evolutionary forces bestow a systematic selective advantage on democracies over autocracies. To the extent that such selective advantages exist, it is essential to understand them in order to assess the future potential of democracy and in order to understand the limits and opportunities within which agents pursuing a democratic agenda are acting.

An evolutionary dynamic is present when processes are directed, behaving as if they were centrally coordinated, when in fact such coordination does not exist. Markets behave in this way and so does the long-term evolution of democracy. But why do levels of democracy evolve over time? Higher levels of democracy emerge as higher levels of action resources become available to the average person in a society. In an era of mass politics, democracies enjoy three distinct selective advantages over autocracies. First, there is a selective advantage by international confrontations. States have been involved in international conflicts and wars, and often the winning states' political regimes replaced the losing states' ones. Success in international confrontations has been related to the type of regime. Democracies usually won the wars they were engaged in, partly because in the long run, they could mobilize their people and resources more effectively. Moreover, democracies tend not to fight each other, avoiding extinguishing their own kind. Autocracies do not have this tendency.

Second, there is a selective advantage by economic performance. Democracies have emerged and persisted in technologically and economically more advanced and powerful states, which partly explains their superiority in international confrontations with autocracies. Democracies have been established in more prosperous economies from the start. In addition, democracies continued to outperform autocracies economically, greatly increasing their initial prosperity advantage over time. Equally important, autocracies repeatedly lost their more prosperous members to the democratic camp.

The third selective advantage of democracies is an advantage by popular support, which is a truly selective force. Because they grant power to the people and because their rulers are selected by the population, democracies tend to have more popular support than autocracies, which enables them to mobilize their people more effectively. This also allows them to limit mass disaffection more effectively than autocracies. Even autocracies that seem stable on the surface, lacking obvious signs of mass opposition, are vulnerable to the "element of surprise" that becomes apparent in democratic revolutions when massive opposition to a regime suddenly emerges and persists, toppling a regime that may have lasted for decades. Democracies are less vulnerable to extinction by popular revolutions. They simply change their rulers through elections.

Democratization and the Future of Politics

The selective advantages of democracy are of such a long-term nature and so deeply rooted in basic developmental processes that there is no reason to assume that the odds will fundamentally turn against democracy in the foreseeable future. Setbacks will occur in specific countries, but the achievements of the global wave of democracy are unlikely to be reversed. But this does not mean that there are no future challenges. Instead, we see a number of challenges on the democratic agenda, which can be formulated in terms of the following questions:

1. Will democracy continue to spread geographically?
2. Will the deficiencies of new democracies, such as those in the former Soviet Union, be overcome?
3. Will the democratic qualities of established democracies be further deepened?

One might also question the viability of the democratic principle in an era in which the major organizational frame of democracy, the nation-state, is said to have lost its significance. And one might question the viability of the democratic principle in a world in which decisive ecological measures seem to be unpopular, though they may be necessary to save our planet.

Spreading Democracy to New Regions

Two important geographical areas have, so far, proved to be relatively immune to the democratic trend: China, and the predominantly Islamic Middle East and North Africa. Anchoring democracy in these two areas would without doubt constitute a major breakthrough for the democratic principle. As far as the Middle East and North Africa are concerned, a sweeping democratic trend throughout the region does not seem likely in the near future. The terror and violence nurtured by the Israeli–Palestinian conflict, Islamic fundamentalism, and the predominance of patrimonial states based on oil rents all amount to powerful obstacles to democratization. In addition, there is throughout much of the Islamic world, but especially in the Middle East, a cultural self-appraisal of Islam as the West's countercivilization—an understanding that is sometimes mirrored in Western views of Islam as its countercivilization. On this basis, democracy is considered to be a Western product in much of the Islamic world, which might disqualify it in the eyes of many people. Evidence from the World Values Surveys indicates that even among those segments of Islamic populations that overtly support democracy, there is often a fundamental misunderstanding of democratic principles. Evidence from the World Values Surveys also suggests that patriarchal-authoritarian values, which are incompatible with democracy, are prevalent in much of the region, particularly the Arab-speaking countries. Although these factors hinder the emergence of democracy, the general idea of democracy is widely approved, and democratic institutions seem to be taking root in some historically Islamic societies such as Turkey.

Consolidating and Improving New Democracies

Many new democracies in Latin America, sub-Saharan Africa, and Central and Eastern Europe show serious deficiencies concerning the rule of law, accountability, and transparency. Not surprisingly then, there is widespread popular cynicism about the integrity of representatives, the trustworthiness of institutions, and policy performance in these new democracies. This popular cynicism often leads to political apathy rather than mass political activism, weakening civil society and placing corrupt leaders under little popular pressure to behave more responsively. But in those new democracies where cynical citizens become "critical citizens" who sustain a high level of elite-challenging mass activities, government is consistently more effective, transparent, and accountable. Civic action matters: Both within new and old democracies, relatively widespread civic action helps in increasing accountable governance. This insight is important. It shows that the quality of democracy is not solely the concern of the ruling elite. It is also, and very markedly so, a matter that concerns the citizens. When they are motivated to put the ruling elite under popular pressure and actually do so, they can improve the quality and effectiveness of governance. There is no reason for civic defeatism.

Deepening Old Democracies

The most obvious aspect of the global democratic trend is the geographical spread of democracy. But the global democratic trend has a second, often forgotten aspect: the deepening of democracy. This occurs even where democracy has been in place for many decades. This trend is well documented in the book *Democracy Transformed? Expanding Political Opportunities in Advanced Industrial Democracies* by Bruce E. Cain, Russell J. Dalton, and Susan E. Scarrow (2005), showing that over the past 25 years most postindustrial democracies have widened the elements of direct democracy, opened channels of citizen participation in policy planning, extended the scope of civic rights, and improved accountability to the public. These institutional changes have been accompanied and driven by cultural changes that gave rise to emancipative values and high levels of sustained elite-challenging actions. In fact, a major reason why long-established democracies show high levels of accountable governance is because they are constantly exposed to popular pressure by increasingly critical citizens. This should affect our views of what kind of citizenry is needed to consolidate democracies and keep them flourishing.

In *The Civic Culture*, Gabriel Almond and Sidney Verba (1963) assumed that in order for democracy to flourish, citizen participation should be limited to the institutional channels of representative democracy, focusing on elections and the activities around them. This view was reinforced by Samuel Huntington's (1968) influential work *Political Order in Changing Societies*, contributing to deep-seated suspicions of noninstitutionalized, assertive citizen action. This suspicion is so deeply ingrained in political science that, even today, prevailing concepts of social capital and civil society still focus on institutionally channeled participation, emphasizing membership and participation in formal associations. By contrast, noninstitutionalized forms of assertive citizen action are rarely recognized in prevailing conceptions of civil society. The dominant view of what sort of citizenry makes and keeps countries democratic needs to be revised. Democracy flourishes with an uncomfortable citizenry that makes life difficult for their rulers, exposing them to constant popular pressure. Democracy requires a citizenry who place a high value on democratic freedoms and are capable of struggling for them—to attain them when they are denied and to sustain them when they are challenged.

Unfortunately, such a citizenry cannot be ordered into existence by elite decree, nor can it be crafted by institutions. Its emergence reflects a more basic process of human empowerment and social and cultural modernization through which people acquire the resources and skills to demand responsive government and the values that motivate them to practice democracy. Democratic institutions can be imposed from outside, but if these conditions are absent, it is likely to be a flawed version of democracy if it survives at all. Anchoring democracy is not just about crafting institutions. It is about shaping the development of economic, cultural, and political modernization of a given society.

Christian W. Haerpfer
University of Aberdeen
Aberdeen, United Kingdom

See also Democracy, Quality; Democracy, Theories of, Democratic Consolidation; Democracy, Types of; Hybrid Regimes, Transition

Further Readings

Almond, G., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations.* Princeton, NJ: Princeton University Press.

Berg-Schlosser, D. (Ed.). (2007). *Democratization: The state of the art.* Opladen, Germany: Barbara Budrich.

Cain, B. E., Dalton, R. J., & Scarrow, S. E. (Eds.). (2005). *Democracy transformed? Expanding political opportunities in advanced industrial democracies.* Oxford, UK: Oxford University Press.

Casper, G., & Taylor, M. (1996). *Negotiating democracy: Transitions from authoritarian rule.* Pittsburgh, PA: Pittsburgh University Press.

Diamond, L. (1992). Economic development and democracy reconsidered. In G. Marks & L. Diamond (Eds.), *Reexamining democracy: Essays in honor of Seymour Martin Lipset* (pp. 93–139). Newbury Park, CA: Sage.

Haerpfer, C. W., & Inglehart, R. F., & Bernhagen, P., & Welzel, C. (2009). *Democratization.* Oxford, UK: Oxford University Press.

Huntington, S. (1968). *Political order in changing societies.* New Haven, CT: Yale University Press.

Inglehart, R. F., & Welzel, C. (2005). *Modernization, cultural change, and democracy.* Cambridge, UK: Cambridge University Press.

Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America, and Post-Communist Europe.* Baltimore: Johns Hopkins University Press.

Lipset, S. M. (1994). The social requisites of democracy revisited. *American Sociological Review, 59*(1), 1–22.

O'Donnell, G., & Schmitter, P. C. (1986). *Transitions from authoritarian rule: Tentative conclusions about uncertain democracies.* Baltimore: Johns Hopkins University Press.

Whitehead, L. (2003). *Democratization: Theory and experience.* Oxford, UK: Oxford University Press.

Dependency Theory

According to dependency theory, Third World countries in the periphery of the global capitalist system are poor and underdeveloped because they are exploited by the advanced capitalist countries in the core. Their major problem is dependency on the core countries. That is the central claim made by dependency theory, a mode of analysis developed primarily by Latin American scholars in the 1960s and 1970s.

Dependency theory emerged as a reaction against the modernization paradigm, which dominated Western liberal approaches to development during the decades after World War II. Modernization theory argued that Third World countries should be expected to follow the same developmental path as taken earlier by developed countries in the West: a progressive journey from a traditional, preindustrial, agrarian society, toward a modern, industrial, mass-consumption society. Development meant overcoming barriers of preindustrial production, backward institutions, and parochial value systems that impeded the process of growth and modernization.

The theoretical endeavors among modernization theorists concerned identification of the full range of impediments to modernization as well as all factors that promote modernization. A famous economic modernization theory by Walt Whitman Rostow specifically stressed that the "takeoff," the crucial push in moving from traditional toward modern, is characterized by a marked increase in modern sector investment to a minimum of 10% of the gross national product. Another critical element concerns the relationship of Third World countries to the world market. Close market relations with the developed countries are seen to have a positive developmental effect on Third World economies. Foreign trade is viewed as a road to market expansion and further growth of the modern sector. Foreign direct investment in the Third World by transnational corporations (TNCs) brings in the much needed modern technology and production skills.

Early Critiques of Liberal Economic Theory

This liberal understanding of development was subjected to increasing criticism during the 1960s and 1970s. That was partly in reaction to the lack of progress in many Third World countries at the time. While growth rates in the developed world reached unprecedented highs in the postwar decades, many Third World countries had difficulties in getting economic development under way. Their economies refused to "take off." That naturally led to increasing dissatisfaction with modernization theory.

Already in the late 1940s, the Argentinean economist and central bank director Raul Prebisch had criticized liberal economic theory in relation to the Third World. In particular, he turned against David Ricardo's famous notion of comparative advantage according to which countries in the periphery would be better off specializing in production and export of raw materials and agricultural products, leaving advanced industrial production to the core countries. Prebisch argued that production from the periphery was subjected to deteriorating terms of trade compared with products from the advanced countries. Two factors were involved. The first is income elasticity, which states that raw materials and agricultural products are less in demand when incomes increase; consumers turn to nonfood items, and technological advance decreases the demand for raw materials. Second, because of strong labor unions in the advanced countries, productivity increases lead to real wage improvements in the core rather than to lower prices for manufactured products. In effect, comparative advantage does not work for the periphery. The core countries keep the benefits of exchange for themselves.

Emergence of Dependency Theory

Dependency theorists were inspired by Prebisch. They took his critique and developed it further, to a full-scale attack on liberal ideas of modernization, specialization, and comparative advantage. According to Theotonio dos Santos (1970), dependency

> is a situation in which the economy of certain countries is conditioned by the development and expansion of another economy to which the former is subjected. . . . [S]ome countries (the dominant ones) can expand and become self-sustaining while other countries (the dependent ones) can do this only as a reflection of that expansion. (p. 231)

Based on this general notion, scholars primarily from the Latin American school of theory developed their particular theories of dependency. In addition to dos Santos, they include Fernando Henrique Cardoso, Andre Gunder Frank, Ruy Mauro Marini, Osvaldo Sunkel, and several others. In the present context, it is relevant to focus on Frank and Cardoso because they represent the most forceful and popular varieties of dependency theory. They are neo-Marxist in the sense that they both draw on classical Marxist analysis of the capital–labor relationship, but they differ from classical Marxism in a basic respect. Unlike Marx, dependency theorists do not expect capitalist development to take root and unfold in the Third World in the same way that capitalism evolved in Western Europe and North America. At the same time, they disagree about the prospects for development in the Third World. Radical theorists, such as Frank, believe that a separation from capitalism and the creation of a popular-based socialist model is the only way forward. Moderate theorists such as Cardoso argue that some form of dependent development can take place in the periphery in context of the capitalist system. In both cases, however, focus is less on future strategies for development and more on present weaknesses of the capitalist system.

Frank was born in Berlin and educated as an economist in Chicago before he came to Chile and Mexico to take up the analysis of dependency. His book from 1967, *Capitalism and Underdevelopment in Latin America*, was a central factor in the establishment of dependency theory as a major approach to the study of the Third World; it was strongly inspired by Paul Baran's study *The Political Economy of Growth* from 1957.

For modernization theorists, "traditional society" was the place where all countries started their process of development and modernization. Frank rejects that view. The starting point for him is not tradition—it is underdevelopment. Underdevelopment is not a condition that once characterized all countries. It is a process within the framework of the global capitalist system to which Third World countries have been subjected: They have been underdeveloped as an intentional by-product of the development of the advanced core countries. Earlier forms of society in the Third World may have been *un*developed; *under*development begins only with the arrival of a global capitalist system. That is, global capitalism in one single process generates development and wealth (in the industrialized world) and underdevelopment and poverty (in the Third World).

This is due to what Frank (1967) calls the three basic contradictions in capitalism: (1) the expropriation/appropriation of economic surplus, (2) the polarization into metropolitan center and peripheral satellites, and (3) the continuity of the capitalist structure and its generation of underdevelopment. In other words, the core countries began the exploitation of the periphery early—in the 16th century in Chile's case. The core appropriates economic surplus from the satellites and uses it for its own development. Some industrial development in the periphery may be possible, but it is dominated by external economic interests; these external forces result in crippled and distorted societal structures inside peripheral countries. To overcome underdevelopment, a delinking from external dominance is required.

Frank's analysis is simple and clear. It is also conceptually imprecise. "Capitalism" is said to have existed since the 16th century; "economic surplus" has been appropriated by the core for several centuries; and "underdevelopment" characterizes the periphery irrespective of the level of industrialization. Consider the example of Brazil. From the sugar mills run by slaves in the 17th century to the advanced petrochemical industries of today where transnational corporations participate, history merely repeats itself, in that global capitalism creates underdevelopment because of the system's basic contradictions. In one sense, Frank's dependency theory turns modernization theory on its head: What comes out of the core is always bad (whereas for modernization theory, it is always good), and capitalism perpetuates underdevelopment. As Aidan Foster-Carter (1976) notes, where modernization theory had difficulties in explaining why "traditional" societies would inevitably move toward modernity, dependency theory has difficulties in explaining why the result must always be underdevelopment rather than development.

The Egyptian economist Samir Amin basically agrees with Frank, but he proposes an analysis that tackles some of the problems mentioned here. Amin focuses on Africa. He analyses the different historical phases of African dependency and argues that the specific structure of peripheral economies is based on export of raw materials and import of (mostly luxury) consumer goods. Yet Amin agrees with Frank that any economic development of the periphery is blocked due to the subordination of the periphery to the core.

Criticism of the "Blocked Development" Theory

From early on, this radical claim of "blocked development" in the periphery had been met with skepticism by several Latin American scholars. Already in 1969, Cardoso and Enzo Faletto had published a book on dependency and development in Latin America rejecting the soundness of a general theory of "blocked development" and arguing in favor of concrete historical analysis of specific situations of dependency. In the mid-1970s, Cardoso published a critique of the radical claims, exposing what he called "five false theses" about capitalism in Latin America. Cardoso argued that some form of "dependent-associated development" was possible in the periphery through a form of capital accumulation similar to capitalism in the advanced countries.

This, then, was the major issue of contention among dependency theorists in the 1970s: Is the periphery doomed to nondevelopment? Or, is it more relevant to speak of a process of "dependent development" formed by the specific national conditions in each country, especially by the relative strength and alliance of social classes and factions? The case of Brazil would appear to support the latter view: Industrial development in that country was shaped by an alliance between elements of the international bourgeoisie, the local bourgeoisie, and public and entrepreneurial bureaucratic elites. The North American sociologist Peter Evans wrote a well-known analysis of Brazil in 1979 based on this view.

Most of the critique of dependency theory, which emerged in the 1970s and 1980s, was directed against the "blocked development" view of Frank, Amin, and others. A number of countries in Southeast Asia, most notably the "Four Tigers" (South Korea, Taiwan, Singapore, and Hong Kong), experienced rapid economic growth combined with world market integration. That was a blow to radical dependency theory's prediction of stagnation and misery in the periphery. The "dependent development" view was less easy to reject. At the same time, Cardoso's theoretical views were also less clear because they tied the fate of peripheral countries to particular historical circumstances. That is a difficult starting point for attempts at a general theory of development and underdevelopment. (Later, Cardoso entered active politics and was president of Brazil for two periods—between 1995 and 2003.)

World Systems Theory

Some of the core claims of radical dependency theory were repeated in the world system analysis by Immanuel Wallerstein. At the same time, Wallerstein repaired some of the theory's weaknesses. His focus was on the capitalist world economy established in "the long 16th century" (1450–1640). It was based on an international division of labor that covered Europe first but soon expanded to the Western hemisphere and later to other parts of the world as well. Within this division of labor, a process of specialization took place; at first, this somehow happened accidentally. For a number of reasons, Northwest Europe was in a better position to diversify its agriculture and to connect it with industrial advances in textiles and shipping. So the capitalist world economy is built on a hierarchy of core areas, peripheral areas, and semiperipheral areas. The core areas contain the advanced and complex economic activities (mass market industries and sophisticated agriculture). Furthermore, these activities are controlled by an indigenous bourgeoisie. Peripheral areas are at the bottom of the hierarchy. They produce staple goods such as grain, wood, and sugar. They often employ slavery or coerced labor. Semiperipheral areas are economically mixed and constitute a middle layer between the upper stratum of the core and the lower stratum of the peripheral countries.

A basic mechanism of the capitalist world economy is unequal exchange. Economic surplus is transferred from the periphery to the core. This transfer is further accentuated by the emergence of strong state machineries in the core and weak state machineries in the periphery. Strong states can enforce unequal exchange on weak ones. In the process of unequal exchange, tensions are created in the system. The semiperiphery has an important function in this regard. It provides an element of political stability because the core countries are not facing a unified opposition. The semiperiphery acts as a buffer or shock absorber. At the same time, the world economy is not entirely static; any single area of the system may change from periphery to semiperiphery, from semiperiphery to core, and vice versa. But Wallerstein emphasizes that the capitalist system as such does not change. It remains a hierarchy of core, semiperiphery, and periphery, characterized by unequal exchange.

Wallerstein sees the end of the Cold War and the destruction of the Soviet bloc as a consequence of the development of the capitalist world economy. However, the long-term prospect is the demise of the capitalist system because the contradictions of that system are now unleashed on a world scale. Success, not failure, is the real threat to global capitalism. When the possibilities for expansion are all used up, the never-ending quest for more profit will lead to new crises in the world capitalist economy, which sooner or later will spell its transformation.

Wallerstein injects some dynamic elements into the picture of unequal exchange and exploitation of the periphery by the core. Single countries can move up and down in the system. The semiperiphery is categorized as a new "buffer zone." Politics and conditions in single countries play a larger role. But still, focus remains on economics and the subordination of the periphery. Recent analyses such as that of John Darwin (2007) reject that view as one dimensional and reductionist. Already by the early 1990s, several areas in the Third World experienced dynamics of development difficult to reconcile with radical dependency analysis even in its more sophisticated Wallerstein version. Instead, more emphasis was laid on the moderate "dependent development" view, which was now combined with new analyses of the internationalization of capital or economic globalization.

New Approaches

Cardoso and Faletto studied the internationalization of capital as seen from the periphery. A new generation of analyses took up this topic as seen from the core countries. This latter approach had already been used by the classical theories of imperialism, for example, those of Lenin, Rosa Luxembourg, John Hobson, and others. The more recent generation of studies focuses on the so-called new international division of labor where industrial production is relocated from the core to the periphery in search of cheap labor and other favorable conditions of operations and/or in search of local markets that could not easily be accessed from the outside. A pioneering study from the three German scholars Folker Fröbel, Jürgen Heinrichs, and Otto Kreye argued that already during the 1960s, a world market for industrial activity emerged. It was based on three factors: (1) a huge supply of cheap labor with basic skills in the Third World and in Eastern Europe, (2) communication and transport technologies that facilitate relocation on a global scale, and (3) production technologies allowing the separation of complex processes of production into relatively simple component parts. The analysis points out that not only labor-intensive industries, such as textiles, clothing, footwear, and toys, but also industries that are associated with the production of electronic parts, machine building, and computer manufacturing are good candidates for global scale manufacturing.

Other studies focused on the attempt to access lucrative domestic markets in the Third World as the major push to internationalization. In Brazil, foreign investment helped create a substantial automobile industry and also participated in the development of petrochemical, computer, and aircraft industries. These developments raised the issue anew of whether foreign capital should be seen as a primarily "constructive" or primarily "destructive" factor in terms of economic development in the periphery. Cardoso had hinted at both: There was development, but there was also dependency. The new generation of analyses further developed this view. It is argued that foreign transnationals can be both "constructive" and "destructive" in the periphery. Whether they are primarily one or the other depends on the conditions under which the companies operate in the host countries. A decisive element in this regard concerns the power position (economically and politically) of local capital relative to that of foreign capital. The stronger the position held by local capital in an industry, the better the possibilities of benefiting from the "constructive" effects and avoiding the "destructive" effects of transnationals (Georg Sørensen, 1983). In other words, successful modernizers, such as the Asian "tigers," China, Brazil, and India, will be better positioned to reap benefits from foreign investment than will very poor and weak states (in sub-Saharan Africa and elsewhere) where local capital is next to nonexistent, and the state is incapable of effective regulation and control.

In sum, even if one can still find examples of radical dependency analysis, it has gone into decline. The 1980s and 1990s saw a strong revival of liberal economic ideas in development thinking. Ronald Reagan's presidency in the United States and Margaret Thatcher's administration in the United Kingdom both promoted liberal policies that emphasized the role of free market forces and the downsizing of state bureaucracies and state regulations. Third World countries were encouraged to pursue similar policies.

Some scholars continue to employ radical dependency ideas in their analysis of macrohistorical developments of capitalism, but most analyses of Third World countries have moved in different directions. The period dominated by liberal economics in the 1980s and part of the 1990s also had dissenting voices. Some of these voices were inspired by the moderate "dependent development" position in their attempt to evaluate prospects for Third World development. This current view is critical of neoliberal free market ideology. We may call it modern mercantilism. Modern mercantilism suggests that Third World countries need to steer between the extremes of economic delinking on the one hand and full integration in the world economy on the other. They argue that the success of economic development of East Asian countries is precisely due to their pursuit of a modern mercantilist strategy. A second core area of development where the modern mercantilists strike a balance concerns the market and the state. Economic liberals argue that free market forces and a minimal role for the state are best for the promotion of economic development. Mercantilists reply that there may be serious flaws in the alleged efficiency of the market. Furthermore, there is no firm support in economic theory for maintaining that state intervention is by definition counterproductive. The fact that some interventions are flawed does not constitute a case against intervention per se. South Korea's and Taiwan's development achievements have been based on states actively working toward building desire structures of production.

Modern mercantilism, then, is based on three ideas: (1) strike a balance between national autonomy and international integration—that is, between incorporation into the world market and self-reliance; (2) strike a balance between state and market—that is, between free market forces and state regulation; (3) foreign direct investment by TNCs can be a strong force for modernization but only provided that TNCs are counterbalanced by local industry and host government supervision.

In conclusion, radical dependency theory is no longer a major approach to Third World development problems, but elements of moderate dependency theory remain centrally relevant for thinking about economic development in the periphery.

Georg Sørensen
Aarhus University
Aarhus, Denmark

See also Imperialism; International Trade; Modernization Theory; Multinational Corporations (MNCs)

Further Readings

Amin, S. (1976). *Unequal development*. Sussex, UK: Harvester Press.

Cardoso, F. H. (1976). *Current theses on Latin American development and dependency: A critique*. Unpublished paper. Bergen University, Norway.

Cardoso, F. H., & Faletto, E. (1979). *Dependency and development in Latin America*. Berkeley: University of California Press. (Original work published 1969)

Darwin, J. (2007). *After Tamerlane: The rise and fall of global empires, 1400–2000*. London: Allen Lane.

dos Santos, T. (1970, May). The structure of dependence. *American Economic Review, 60,* 231–236.

Foster-Carter, A. (1976). From Rostow to Gunder Frank: Conflicting paradigms in the analysis of underdevelopment. *World Development, 4*(3), 167–180.

Frank, A. G. (1967). *Capitalism and underdevelopment in Latin America*. New York: Monthly Review Press.

Kay, C. (1989). *Latin American theories of development and underdevelopment*. London: Routledge.

Sørensen, G. (1983). *Transnational corporations in peripheral societies*. Aalborg, Denmark: Aalborg University Press.

Wallerstein, I. (1974). *The modern world system, I*. New York: Academic Press.

Deregulation

In the broadest sense, deregulation can be described as the nonmarginal reduction or total withdrawal of the state's legal control over the social and economic activities of firms and citizens. Although deregulation is often associated with the aim of increasing competition in economic sectors, such as transport, energy, and audiovisual markets, the concept covers social and administrative regulation too. In this case, competition is not the only aim of deregulation, since social regulation can be inspired by the notion of increasing access (i.e., via simplification) and participation or the aim of respecting individual choice in delicate areas (e.g., the consumption of recreational drugs). Administrative deregulation has triggered ambitious government programs to reduce administrative burdens affecting citizens and firms in Europe, North America, and some developing countries.

However, the notion of deregulation must be put in context to identify its purpose and implications as well as the inextricable link with its counterpart—regulation. Indeed, when new policy is designed, deregulation is one of the instruments of regulatory choice, together with self-regulation, regulation by information, coregulation, command, and control regulation.

Deregulation can be analyzed from both historical and theoretical angles. In turn, these perspectives are linked to three clusters of questions: (1) How, where, and when does deregulation occur? (2) Why does it occur? (3) What (actors or structural forces) caused it? In addition, a geographical angle is also needed, especially in light of debates on regulatory reform and the diverse meaning that deregulation has in different continents, transition economies, and leading economies.

Historical Perspective

Historically, deregulation emerged as a response to the welfare and Keynesian programs of the 1950s and 1960s and the economic instability of the 1970s. Conventionally, deregulation is associated with the rise to power of Margaret Thatcher in the United Kingdom (1979) and Ronald Reagan in the United States (1981). In this received view, deregulation is associated with the economic paradigm of neoliberalism. The international diffusion of deregulation was facilitated by the programs of the International Monetary Fund (IMF; e.g., in Latin America), the ideas cementing the Washington Consensus, and the promotion of approaches to deregulation, such as the ones pursued by labor governments in Australia and New Zealand. When labor governments joined the bandwagon of deregulation, the international spread of the phenomenon seemed to some observers universal and unstoppable. The conventional wisdom, however, ignores that (a) as a matter of fact, the total amount of regulation continued to grow in the 1980s and is still on the rise; (b) the association between deregulation and other so-called neoliberal ideas is dubious, as shown by the spurious correlation between deregulation and privatization; (c) there was strong political opposition to deregulation ideas ("genuine," "persistent," and "vicious," according to Sam Peltzman), hence they were never uncontested and triumphant; with the financial crisis triggered by the credit crunch of 2008–2009, deregulatory ideas lost much of their intellectual appeal, amid several calls for "new" forms of regulatory capitalism; and (d) in continental Europe at least, the deregulation of domestic utilities and other crucial markets was caused by a newly emerging regulatory power, that is, the European Union (EU) as "regulatory state." National rules were removed and remodeled, but this happened in the context of a transfer of regulatory powers to the supranational European level to facilitate the creation of the EU Internal Market.

Theoretical Perspectives

The theoretical perspectives address the question of why this phenomenon occurs in the first place. Basically, there are four theoretical rationales or explanations. The first looks at the politics of ideas. As mentioned, the conventional wisdom of deregulation points in the direction of the ideas of the New Right, the economic policy missionaries of international organizations, and the doctrines of economic rationalism. However, a more accurate interpretation in terms of politics of ideas was provided by Martha Derthick and Paul J. Quirk. They argue that deregulation was the product of successful, convincing, evidence-based economic arguments that demonstrated the large diffuse benefits of deregulation to the public and politicians, thus overcoming the resistance of client politics, that is, the narrow, concentrated benefits from regulatory rents. But interest group theorists propose a second explanation, rooted in the advantages that firms exposed to international competition gain from deregulation (a variant of this explanation looks at the power of corporate consumers or business consumers). A third explanation turns to historical narratives and sees deregulation as the effect of previous dynamics of regulatory policy—a case of policy self-destruction once regulation reached its politically sustainable climax. By focusing on the formal analysis of the rules of the political system (but often blended with historical narratives of specific cycles of deregulation), a fourth explanation sheds light on the institutional capacity for reform, showing when and how institutions, often in combination with the emergence of policy entrepreneurs or reform champions, can overcome the inertial dynamics of regulatory agencies and client politics. This fourth explanation ties in with the question of who (de)regulates. Comparative research has explained the different incentives and constraints to deregulation, how institutional rules and legacies provide opportunity structures to reform agents, and finally how these structures can be changed.

The natural complement to these theoretical interrogations is provided by empirical studies of deregulatory policies and their actual outcome in several sectors. Analyses often focus on the economic deregulation of specific industries and show the redistribution of the costs and benefits of deregulation across different societal groups. While the effects of (economic) deregulation on prices, competition, and market structure are clearly identifiable in several instances, at times, the final outcome of this policy cannot be fully separated from other variables such as technological change and, more broadly, from other ingredients of regulatory reform (e.g., reregulation).

Regulatory Strategies

Contemporary capitalism is an era of regulatory flux where regulatory, deregulatory, and reregulatory shifts are occurring simultaneously. New regulatory domains have emerged, such as risk, the regulation of private security companies, financial (re)regulation, and corporate governance, as well as ambitious attempts for forging international regulatory cooperation in areas such as climate change and intellectual property rights.

Deregulation was widely adopted by developing countries with the aims of increasing foreign direct investment and reducing regulatory burdens on endogenous growth. Yet deregulatory failures, especially but not exclusively in developing countries, have triggered a recalibration of policy strategies. The emphasis has shifted from deregulation to regulatory quality—the question is not the total level of regulation but the efficiency, accountability, consistency, and transparency of regulation. In turn, research on regulatory quality has shown that it cannot be achieved by simply clamping down on the total number of rules. It requires a proper institutional design of regulatory bodies and regulatory oversight institutions. This explains the political attention to institutional innovations such as independent regulatory authorities and other nonmajoritarian bodies, the economic analysis of proposed regulation (also known as regulatory impact assessment), and the role of central oversight units in charge of regulatory quality goals (Giandomenico Majone, 2001; OECD, 2002). It is regulatory quality, rather than the total level of regulation, that has been identified as an important factor of growth (see Hossein Jalilian, Colin Kirkpatrick, & David Parker, 2007, for econometric evidence on developing countries).

Thus, although at the time some authors interpreted deregulation as yet another sign of the end of history, the academic analysis and the policy debate on deregulation is very much alive 30 years after Ronald Reagan famously said on January 20, 1981, in his inauguration speech: "Government is not the solution to our problem, government is the problem."

Claudio M. Radaelli and Lorna Schrefler
University of Exeter
Exeter, United Kingdom

Further Readings

Armstrong, K. A. (2000). *Regulation, deregulation, re-regulation*. London: Kogan Page.

Ayers, I., & Braithwaite, J. (1992). *Responsive regulation: Transcending the deregulation debate*. Oxford, UK: Oxford University Press.

Derthick, M., & Quirk, P. J. (1985). *The politics of deregulation*. Washington, DC: Brookings Institution Press.

Hood, C. (1994). *Explaining economic policy reversals*. Buckingham, UK: Open University Press.

Horwitz, R. B. (1986). Understanding deregulation. In Structures of capital [Special double issue]. *Theory and Society, 15*(1/2), 139–174.

Jalilian, H., Kirkpatrick, C., & Parker, D. (2007). The impact of regulation on economic growth in developing countries: A cross-country analysis. *World Development, 35*(1), 87–103.

Majone, G. (1990). *Deregulation or re-regulation? Regulatory reform in Europe and the United States*. London: Francis Pinter.

Majone, G. D. (2001). Nonmajoritarian institutions and the limits of democratic governance: A political transaction-cost approach. *Journal of Institutional and Theoretical Economics, 157*(1), 57–78.

Organisation for Economic Co-operation and Development. (2002). *Regulatory policies in OECD countries: From interventionism to regulatory governance*. Paris: Author.

Peltzman, S. (1989). The economic theory of regulation after a decade of deregulation. In M. N. Baily & C. Winston (Eds.), *The Brookings papers on economic activity* (pp. 1–48). Washington, DC: Brookings Institutions.

Stigler, G. (1981). Regulation in theory and practice: An overview: Comment. In G. Fromm (Ed.), *Studies in public regulation* (pp. 73–77). Cambridge: MIT Press.
Winston, C. (1993). Economic deregulation: Days of reckoning for microeconomists. *Journal of Economic Literature, 31*(3), 1263–1289.

DÉTENTE

Détente is a word of French origin meaning "a relaxation of tension." In its traditional diplomatic usage, détente has most commonly been used to indicate a reduction of tension between two adversarial states. Détente has alternately been used to define a condition, a process, a policy, and a period of history. The concept is not synonymous with peace or a natural harmony of interests. In fact, a central element to any definition of détente is the precondition of intense tension and rivalry between two states with substantial conflicting interests. Détente, as Richard W. Stevenson (1985) proposed, might be best defined as "the process of easing tension between states whose interests are so radically divergent that reconciliation is inherently limited" (p. 11). Within that process, policies of engagement, accommodation, and cooperation may be pursued by adversarial states to reduce the tension created by their underlying conflict, which left unrestrained might eventually lead to war.

Détente first became part of diplomatic parlance in 1908 as the European great powers attempted to reduce the tension created by the then crises in Morocco and other imperial rivalries. However, détente is most famously associated with the foreign policy of former U.S. President Richard Nixon and National Security Advisor—eventually Secretary of State—Henry Kissinger during the 1970s. When Nixon came to power in 1969, constant threat of nuclear war and deep hostility of the Cold War had characterized U.S.–Soviet relations for more than 2 decades. To be sure, the real fear of nuclear catastrophe created by the Cuban Missile Crisis of 1962 introduced an element of caution into the Cold War era and was perhaps the impetus for agreements such as the Partial Test Ban Treaty and the Nuclear Non-proliferation Treaty, but U.S.–Soviet relations throughout the late 1960s continued to be marked by suspicion and intense ideological rivalry. International conflict, such as the Soviet invasion of Czechoslovakia in 1968 and continuing American war against communist forces in Vietnam, only heightened the danger of U.S.–Soviet military confrontation and possible escalation.

Nixon and Kissinger pursued détente in this strategic context. The first indication of the new U.S. approach to the Soviet Union and China appeared in Nixon's inaugural address on January 20, 1969, when he declared, "After a period of confrontation, we are entering an era of negotiation." Furthermore, Nixon sketched a rough outline of the framework for détente when he invited the adversaries of the United States into a "peaceful competition" and proposed cooperation to "reduce the burden of arms." The elements of arms control and restrained competition would come to be central features of the Nixon–Kissinger détente efforts.

The mechanics of détente were primarily left to Kissinger, who as Nixon's national security advisor centralized control of foreign policy in the White House under the auspices of the National Security Council (NSC). Circumventing the traditional foreign policy bureaucracy of the State Department, Kissinger established a "back channel" with the Soviet ambassador to the United States, Anatoly Dobrynin. This back channel provided an outlet for Nixon to communicate directly with Communist Party General Secretary Leonid Brezhnev on sensitive bilateral issues and during emerging crises. The origins of many of the significant U.S.–Soviet agreements and accords of the détente era can be traced to the Kissinger–Dobrynin back channel.

The Nixon–Kissinger approach to détente can be viewed as an attempt to find a modus vivendi with the ascendant Soviet Union. By the late 1960s, the Soviet Union had achieved a rough parity with the United States in terms of strategic military power, and détente provided Nixon and Kissinger with a method of managing the Soviet Union's rise to true superpower status while concurrently attempting to mitigate the negative aftermath of the Vietnam War on U.S. power and prestige. Nixon and Kissinger promoted the concept of linking multiple policy issues to gain greater diplomatic leverage over the Soviet Union. In their view, linkage politics provided a method of addressing a broad range of issues while trying to prevent any single issue from derailing the overall process and milieu created by détente. Despite the Soviet Union's emergence as a global power in the military sense, Soviet power was not nearly as multidimensional as U.S. power. As a result, there were power asymmetries, which provided opportunities—in Kissinger's view—for the United States to exploit these asymmetries through linkage politics thereby restraining Soviet actions.

Détente between the United States and the Soviet Union, as Mike Bowker and Phil Williams have noted, consisted of four primary and interconnected elements: (1) European normalization, (2) arms control, (3) codes of conduct, and (4) East–West trade. The issue of normalizing European relations with the Soviet Union and Eastern Bloc countries had remained unresolved following the end of World War II, and in 1969, West German Chancellor Willy Brandt instituted a policy of *Ostpolitik* (Eastern politics) designed to improve and increase relations between West Germany, East Germany, and the Soviet Union. Détente in Europe proceeded slightly apart from superpower détente despite Kissinger's efforts to prevent any separation of the two diplomatic tracks. In 1970, the Moscow Treaty formally recognizing Europe's post–World War II borders and the division of Germany was signed by West Germany and the Soviet Union. The historically thorny issue of Berlin involved greater participation by the superpowers, and the Kissinger-Dobrynin back channel proved to be a useful device during these negotiations. The formal agreement on the status of Berlin was signed by the four occupying powers—(1) the United States, (2) the United Kingdom, (3) France, and (4) the former USSR—in September 1971. Several other treaties normalizing the borders and relations between West Germany and other Eastern Communist states, including Poland (1970), East Germany (1972), and Czechoslovakia (1973), were enacted during the early years of European détente. The normalization of European relations through détente did much to relax the tension on what had historically been one of the Cold War's central fronts.

The issue of bilateral arms control was perhaps the centerpiece of détente between the United States and the Soviet Union. Preliminary negotiations on arms control, commonly known as the Strategic Arms Limitation Talks (SALT), began in November 1969, but little progress was made until the Kissinger–Dobrynin back channel reportedly produced a breakthrough in May 1971. Protracted negotiations resulted in the Anti-Ballistic Missile (ABM) Treaty and the Interim Agreement on the Limitation of Strategic Offensive Arms, which were signed in May 1972 by Nixon and Leonid Brezhnev at a highly publicized summit meeting in Moscow. The ABM Treaty limited the number of missile defense sites in either country to a total of two and prohibited the further development, testing, and deployment of various types of ABM capabilities. The Interim Agreement limited the number of land-based Inter-Continental Ballistic Missiles (ICBMs), Submarine Launched Ballistic Missiles (SLBMs), and ballistic missile submarines. The ABM Treaty and the Interim Agreement arising from SALT were hailed as significant diplomatic accomplishments critical for sustaining superpower détente.

The 1972 Moscow summit also produced diplomatic agreements designed to provide a framework for limiting the possibility of escalation resulting from superpower competition. The Basic Principles Agreement (BPA), also known as the Basic Principles of U.S.–Soviet Relations, represented a loose code of conduct for the continuation of détente. The superpowers pledged to avoid actions that might increase bilateral tension and the risk of military confrontation, refrain from threats or use of force against allies or each other, urgently consult each other during crises featuring a possibility of a nuclear conflict, and use negotiation to settle disputes peacefully. Additionally, the Agreement on Avoiding Incidents at Sea was also signed at the Moscow summit. Nixon's desire for "peaceful competition" and Kissinger's plan for using détente to construct a more stable international system based on a triangular—United States, former USSR, and China—balance of power were readily apparent in the form and substance of these agreements. Similarly, the June 1973 U.S.–Soviet summit meeting in Washington, D.C., resulted in further diplomatic agreements based on the emerging détente. Nixon and Brezhnev signed a declaration of principles designed to jumpstart the stalled Strategic Arms Reduction Treaty II (START II) negotiations, and they also signed the Agreement on the Prevention of Nuclear War that elaborated on the Basic Principles Agreement.

Improved U.S.–Soviet relations in the early 1970s also provided an opportunity to expand commercial ties. U.S.–Soviet trade dramatically increased during the early years of détente, and there was even an effort to grant most favored nation (MFN) status to the Soviet Union. However, the element of increasing trade with the Soviet Union was quite controversial in U.S. domestic politics, especially after a massive Soviet purchase of U.S. grain at subsidized prices caused a sharp increase in U.S. domestic grain prices and turned into a political debacle for the Nixon administration.

Détente during this era also faced domestic and international challenges. Whereas Nixon's rapprochement with China was most likely the key factor originally motivating the Soviet Union to pursue détente, other international events undermined the process. U.S. involvement in Vietnam was a constant source of friction, especially since Nixon and Kissinger believed that the Soviet Union could exert greater pressure on North Vietnam to conclude a peace agreement. The 1973 Arab–Israeli War strained détente to its limits by heightening the risks of escalation—the United States increased its nuclear alert at one point in the crisis—and brought into question the efficacy of the BPA in restraining superpower actions. Likewise, the Soviet involvement in Angola in 1975 further weakened détente, led to a cessation of U.S.–Soviet trade, and prompted President Gerald Ford to withdraw support for granting MFN status to the Soviets.

The domestic political challenges to détente in the United States were no less severe. By circumventing the traditional foreign policy bureaucracy and emphasizing secrecy and executive control, the Nixon–Kissinger approach failed to build strong institutional support for détente. This lack of domestic support was compounded by presidential electoral politics and congressional opposition. Presidential politics led Nixon and Kissinger to overstate the diplomatic successes of 1972 and oversell the future promise of détente, and the Watergate scandal only added to Nixon's political difficulties. Congressional opposition, especially from Senator Henry "Scoop" Jackson (D-Washington), further complicated the overall process. The Jackson-Vanik Amendment of 1974 is a case in point. Underlying the entire issue of domestic support for détente was the worldview of the average American citizen. As Paul Lauren, Gordon Craig, and Alexander George (2007) note,

> The pursuit of *détente* required a more sophisticated mental and emotional analysis than that of Cold War rivalry. . . . Members of Congress and the public often were unable or unwilling to grasp the subtleties of a strategy that combined both sticks of punishment and carrots of incentives. (pp. 102–103)

An example of that is provided by the Helsinki Accords of 1975, scoffed at by the Soviets in tandem with elements of the U.S. public but not insignificant in ending the Cold War.

The last vestiges of superpower détente perished in the wake of the 1979 Soviet invasion of Afghanistan. The Carter administration removed the SALT II Treaty from Senate consideration, increased military spending, and initiated the reevaluation of U.S. nuclear war-fighting strategy. With the election of Reagan in 1980, the Cold War returned in earnest until the Soviet Union's eventual collapse in 1991.

Matthew Gratias and Patrick James
University of Southern California
Los Angeles, California, United States

See also Arms Race; Diplomacy; Disarmament; Superpower

Further Readings

Bowker, M., & Williams, P. (1988). *Superpower détente: A reappraisal*. Newbury Park, CA: Sage.

Garthoff, R. L. (1985). *Détente and confrontation: American-Soviet relations from Nixon to Reagan*. Washington, DC: Brookings Institution Press.

Keefer, E. C., Geyer, D. C., Selvage, D. E., Kissinger, H., & Dobrynin, A. F. (2007). *Soviet-American relations: The détente years, 1969–1972* (Department of State Publication No. 11438). Washington, DC: U.S. GPO.

Lauren, P. G., Craig, G. A., & George, A. L. (2007). *Force and statecraft: Diplomatic challenges of our time* (4th ed.). Oxford, UK: Oxford University Press.

Litwak, R. S. (1984). *Détente and the Nixon Doctrine: American foreign policy and the pursuit of stability, 1969–1976*. Cambridge, UK: Cambridge University Press.

Stevenson, R. W. (1985). *The rise and fall of détente: Relaxations of tension in US-Soviet relations, 1953–84*. Urbana: University of Illinois Press.

Deterrence

See Arms Race

Developing World and International Relations

Cleavages matter in all social relations. Given that international relations are a specific form of social relations, it is normal that the role of cleavages is taken seriously. Cleavages can be geographic, political, cultural, or economic. They indicate the extent of polarization of the world around different criteria. For instance, cultural cleavages derive from various differences, including religion, ethnicity, and nationality. Political cleavages are linked to differences in ideology or institutional setup, such as authoritarian regimes versus democratic ones. Economic cleavages have to do with the level of development. In international relations, they are translated through the antagonism between the developed world (the Global North) and the developing world (the Global South). The concept of the developing world is an important one that has acquired an enduring meaning in international relations. This entry examines the concept of the developing world through a discussion of the cleavage between the North and the South.

The developing world cannot be understood without reference to the developed world—they form an antagonistic couple. Each of these worlds has specific attributes. It could be said that while the developed world is a world of high economic performance and standard of living, the developing world is a world of relative or absolute deprivation. However, it is important to bear in mind that the concept of the developing world is an overgeneralization of a complex reality. Instead of a single developing world, there are several such worlds. The concept of the developing world does not refer to the same reality in Asia and in Africa. The level of poverty and the lack of industrialization are not the same. For example, in 1970, China was viewed as an underdeveloped country in relation to the average Northern developed country, not underdeveloped African countries. There is a hierarchy within the developed world in which some countries of the South are seen as nearer to those of the North. Moreover, the developed world is not a monolithic world. This concept of development is dynamic: One country can be considered part of the developing world at one time and of the developed world at another. For example, today, China is viewed as much more economically powerful than it was several decades ago. While some Arab countries such as Libya, Saudi Arabia, and Kuwait are global financial players due to the incomes generated by oil, they remain poorly industrialized and have a lower literacy rate than other developed countries. Within the developing world, countries can be classified as having relatively low, medium, or high levels of development. Therefore, it appears that the concept of developing world is elastic, and it can easily be simplified when those realities are under scrutiny. Thus, the notion of developing world should be considered as relative and provisional. Through the mediation of international institutions, the dynamics of international relations has progressively defined the notion of developing world. For instance, the United Nations Development Programme (UNDP) uses its Human Development Index to measure development in terms of several indicators such as life expectancy at birth, adult literacy rate, gross domestic product (GDP) per capita, and human poverty. The UNDP's Millennium Development Goals are another entryway to understanding the developing world. The goals, adopted in 2000 by many members of the international community, identify eight objectives to be achieved by 2015. For example, the first goal, "to eradicate extreme poverty and hunger," includes the target of reducing by one half the proportion of people living on less than $1 per day. The UNDP goals are a way of identifying the objective reality that characterizes developing countries. However, a country's recognition and identification of itself as underdeveloped also constitutes a subjective reality; the underdevelopment status is also created and demanded by the underdeveloped countries. A supplementary illustration of this point is an African country that seeks World Bank status of "highly and poorly indebted country" to receive international relief.

From the preceding discussion, it appears that the developing world is an element in international relations. Here, two meanings of international relations are to be taken into consideration: first, the phenomenon of international relations in which the "developing world" intervenes as an actor and, second, the concept of international relations as a science that has to explain the dynamics of the "developing world."

The Phenomenon of International Relations and the Developing World

Like all social relations, international relations are structured by the balance of forces. In this context, power is based on the level of development. The developed world is the dominant world, the world of the powerful, while the developing world is the dominated world, the world of the powerless. Thus, the developing world stands as testimony to the inequality in international relations and constitutes a specific field of action.

The existence of an international hierarchy is confirmed by the category of "developing world." It is the world that, compared with the one that is developed, is considered backward. In contemporary international relations, the developing world has often been named the Third World. This terminology comes from the period during the French Revolution, when *Third Estate* referred to the commoners. According to this analogy, the term *Third World* was used for countries that were neither Western capitalist countries (First World) nor the socialist countries of Eastern and Central Europe (Second World). The term *Third World* indicates the lower position occupied by countries of Africa, Asia, and Latin America. Despite the collapse of Marxism and the Second World, the notion of Third World is still used to designate the developing world. It is the world on which developed countries wield their power through various means: economic, cultural, political, or military.

Economically, the lower position is reflected by the weakness of the national currencies, raw materials dependent on the trade structure, high rate of human poverty, or insufficient industrialization. These elements, with variations, are common to developing countries and explain their dependence on the developed world. The inequality of developing countries is well captured by the Millennium Development Goals: eradicate extreme poverty and hunger, achieve universal primary education, promote gender equality and empower women, reduce child mortality, and improve maternal health. The fact that these objectives are proposed as goals is an indication of the insufficiencies that characterize the developing world. Therefore, in international relations, developing countries are to a large extent subordinate to developed countries, which create explicitly or implicitly international clientelism: Patron states are developed states; client states are developing states. Clientelism makes it difficult for developing states to have diplomatic autonomy or to go beyond vocal affirmation of sovereignty. However, the economic dependence of the developing world is not an absolute. Ideological, political, or religious parameters may determine the diplomatic demarche of a developing state. For example, the overreliance of Egypt on U.S. economic aid has not led to Egypt aligning with the United States on issues regarding the Middle East.

As a specific space of action, the "developing world" has at least two meanings. The "developing world" emerges from the solidarity approach to international relations used by states and international organizations in their action. Each developed state defines its south. For instance, while the United States had privileged South Asia and Latin America during the Cold War era, France was organizing its aid policy with an emphasis on Africa. International organizations such as the World Bank or UNDP elaborate quantitative criteria to measure underdevelopment and then rationalize the contribution of the international community. In this regard, the "developing world" is the space filled by world generosity. The "developing world" has influenced the structuring of international relations through the agencies, funds, and programs aimed at the development of Third World countries. Inasmuch as the "developing world" induces the creation of development organizations, international institutions crystallize the existence of the "developing world."

The "developing world" has also emerged as a distinct language of international relations. The "developing world" was the world politically created by peoples of Africa and Asia meeting in Bandung in 1955; the Bandung conference was the beginning of the collective claim of Third World countries for decolonization and development. From this conference, a collective consciousness of belonging to the underdeveloped world and a clear resolution to change international relations arose. For this reason, the trademark of developing countries in international relations has been (and still is) the denunciation of the international order. In the same way that Westphalia constitutes a mythical reference for the foundation of European public order, Bandung is a reference in international relations. Its revolutionary dimension should not be minimized. Even though notions of equality and noninterference in internal affairs were already part of the language of international relations, these notions were applied only within the European world. They did not guide relations with non-European states. That is why European public law has not outlawed colonization or imperialism. The Bandung conference was a Third World appropriation of principles of European public law and the United Nations Charter. The creation of the Non-Aligned Movement in 1961 in Belgrade was part of the same dynamics. The developing world is a world in rebellion against the international order: the Economic International Order and the Information and Communication World Order.

The clash between the "developing world" and the "developed world" is relative. There are alliances that transcend cleavages. The French aid policy toward its former colonies of sub-Saharan Africa had been an instrument of political control. In return, a small developing country such as Gabon has influenced France through oil exportation. The Non-Aligned Movement as a way to institutionalize the ideological autonomy of developing countries was counterbalanced by the Marxist orientations of its main members such as Cuba and Ghana. Therefore, the "developing world" does not refer to an autonomous entity. There is even internal rivalry due to external interference or local stakes. The rivalry between the presidents of Côte d'Ivoire and Guinea during the 1960s cannot be explained only by the dependence of Felix Houphouet-Boigny of Cote d'Ivoire on France and the rebellion of Sekou Touré of Guinea. The regional leadership in French-speaking Africa, which concerns states and individuals, was also at stake. In summary, economic cleavages are not sufficient conditions for the formation of two fighting communities: "them" and "us."

However, to some extent, developing countries have succeeded in constructing the "North" as the enemy of the "South" through a mobilization of the history of colonization, the systematization of neocolonialism, and the explanation of underdevelopment by exogenous factors. Thus, the recognition of the fluidity of each pole and the relativity of the cleavage between the "North" and the "South" should not lead to the denial of this opposition. If a misunderstanding does exist, it is a workable one. In international relations, the developing world has been institutionalized as a category of vision, division, and action. It is a basis of solidarities of both resemblance and interests: solidarity of resemblance through the regrouping of developing states in circles such as the Group of 77 and solidarity of interests due to the relative similarity of the economic problems of developing states, as exemplified in the dialogue between the European Union (EU) and the African, Caribbean, and Pacific Group of States. It is not just chance that despite the differences that exist between the developing states of Latin America, their respective national and regional diplomatic agendas are dominated by the theme of international democratization, which is in fact a euphemism for the decolonization of international relations.

The Concept of International Relations and the Developing World

The concept of international relations refers to the science of the specific social relations characterized as international relations. As a body of organized knowledge, the science of international relations does not explain the situation and the role of the states in the same way. The science of international relations does not just involve general knowledge but also particular knowledge of given situations. Therefore, the science of international relations as related to the developing world involves tension between universalism and relativism. As with every science, international relations has general concepts such as those of the state, sovereignty, and power that are applied to the whole world regardless of the level of development. States in the North and the South are both discussed with territory, population, and government as bases. Such a generalization is a lazy one. Territory does not have the same meaning in different political and cultural contexts. The dialectics of unique and multiple meaning(s) are useful. The globalization of the state is paired with its vernacularization. Therefore, its understanding has to be both on global and local levels. A general science of international relations is to some extent an epistemological obstacle for the study of the developing world, because it blurs specificities and differences. It is in this context of local knowledge that a fertile science of international relations is a particular science, that is, a science that opts for the relativity of general categories and concepts. In this regard, it must be acknowledged that some paradigms of international relations possess a real heuristic power to study the role and state of the developing world. However, while universal concepts may be useful in some contexts, it is also important to recognize that seemingly straightforward concepts such as state and sovereignty can mask inequalities and impede the development of knowledge about the developing world, which requires local as well as global knowledge.

The status of the underdeveloped world in international relations gave birth to a specific theorization, and under this rubric some theories of development and underdevelopment, as Ronald H. Chilcote (1981) pointed out, can be located. The dominant explanatory paradigm of the developing world in international relations is Marxist oriented even though societies in Africa, Asia, and Latin America were not the main concern of Karl Marx's thought; the developing world is an interesting terrain for critical international relations theories and more precisely for the use of Marx's writings. Three main themes are taken into consideration by the Marxist-oriented approach: (1) underdevelopment, (2) dependency, and (3) imperialism.

According to the interpretation of the Marxist-oriented approach presented by Samir Amin (1974, 1976), Walter Rodney (1972), and Andre G. Frank (1966), the formation of an underdeveloped world can only be understood in the context of international relations where capitalism evolves as a process of relationship between unequal partners. The underdeveloped world is the product of the development of the capitalist mode of production at the periphery. Therefore, the relations between developed countries and underdeveloped ones are like metropole–satellite relations. This school of thought perceives underdevelopment as related to capitalist, colonialist, and imperialist exploitation. From this perspective, the notion of the developing world is meaningless; only that of an underdeveloped world matters, because Third World countries cannot be liberated from their backward status. Underdeveloped countries and developed ones are perceived as completely different in international relations: The development is unequal because underdeveloped countries are bound to organize their economies according to the needs of the dominant ones.

Dependency of peripheral states on central ones is one of the main characteristics of the international relations structure (Fernando Enrique Cardoso, 1973). The argument of the dependence of the underdeveloped world is based on several elements among which are internal colonialism, which is the domination of the economy by the metropolis, and the extraversion toward central states of the economy.

Imperialism is proposed as an explanation for the actions of some dominant nations in the world. The independence of the underdeveloped states is seen as formal. The actual reality, however, involves the concept of neocolonialism. Inspired by Vladimir Lenin, who considered imperialism to be the highest stage of capitalism, Kwame Nkrumah assumed neocolonialism to be the supreme stage of imperialism.

As briefly summarized, the Marxist-oriented approach to international relations positions the developing world in a specific way in international relations. Contrary to positivist approaches such as realism, which take the position of underdeveloped states as given, the Marxist-oriented approach historicizes the weakness of those states: Where does underdevelopment come from? Under this theoretical prism, the classical concepts of states and sovereignty appear as variables. States and sovereignty do not have the same meaning in metropolitan states (colonizers) and satellite states (the colonies). It is clear that the sovereignty of the latter is more normative than that of the former. However, the Marxist critique of international relations is to be nuanced in light of the experience of the underdeveloped states. First, the fiction of sovereignty and equality among states, be they from the North or the South, has led to a positive reality of decolonization (often not the product of the colonial power's goodwill) and of the ability of weak southern states to oppose strong northern states through diplomatic engagement made possible by the legal and formal superstructure of international relations that can affect the balance of forces. Second, the excessive focus on the dependence dimension of the relationship between the North and the South hides the complex reality of the interdependence that leads to compromise or a variable sum game. Third, the presentation of the underdeveloped world as clay in the hands of the developed world is far from the reality of the relative autonomy of Third World leaders and their capacity to contract alliances according to their interests.

The classical realist paradigm of international relations has indirect consequences in the explanation of the underdeveloped world. It is known that, historically, realism is elaborated from the experiences of major Western states. It has not been easy for classical realism to apprehend the underdeveloped states as they are. During the Cold War period, these states were seen under the prism and of their relations with the two superpowers. Their relative autonomy was ignored. However, under specific conditions, realism can help understand the underdeveloped world in international relations.

Classical realism demystifies the idea of the "underdeveloped world" as an actor in international relations. The unit of analysis and level of observation are constituted by the "state." The notion of "underdeveloped world" is de-globalized for the benefit of underdeveloped states. These states are recognized as being different, with divergent or convergent national interests and with peaceful or conflict-ridden relations. Based on the universal concepts of state, power, and sovereignty, classical realism can then contribute to the de-provincialization of underdeveloped states in international relations. They are states like others. This does not mean that classical realism disregards absolutely sets of states and recognizes only peculiar states. Sets of states are considered when they are political, economic, or military alliances, for instance; an alliance with the underdeveloped world is not an alliance to be taken for granted: There are often alliances between Northern states and Southern states or among some Southern states that crystallize the heterogeneity of the underdeveloped world. At the same time that realism prevents the illusion of spontaneous formation of an alliance, it destroys the illusion of the similarity of underdeveloped states: Like other states, underdeveloped states have distinct resources. Those resources determine their power position. Among underdeveloped states, there are some states that are weaker than the others. Therefore, the realist image of a hierarchy of states is still meaningful.

The redefinition of realism under the traits of "ethnorealism" (Paul Roe, 1999) or "subaltern realism" (Mohammed Ayoob, 1998) constitutes operating attempts to adapt realism to the specificity of Third World states. Both adaptations of classical realism aim at shedding more light on the intrastate dimension of international relations. Linked to the experience of Western states where war and peace are supposed to depend on interstate relations, classical realism does not understand the dynamics of internal conflicts that are preponderant in the underdeveloped world. To cross the boundary that stops classical realism from analyzing the intrastate level, ethnorealism and subaltern realism were proposed. The conceptual basis of international relations is to be reinvented from this perspective. Initially, the great divide is legitimated by the belief that "inside" is the place of order—peace—while the "outside" is dominated by disorder. In several states of the underdeveloped world, internal life is characterized by the intensification of conflicts. In Africa, there are more intrastate wars than interstate ones (Luc Sindjoun, 2002). Thus, it could be said that the study of the underdeveloped world implies a reformulation of classical realism.

The situation of the underdeveloped world in international relations has been a pretext for the denunciation of international relations theory. In fact, this is based on the relationship that is supposed to exist between indigenous culture and Western domination (Homi Bhabha, 2007). An epistemology of the dominant international relations knowledge shows that there is an interaction between realism and imperialism. That is the reason for the call for a postcolonial theory of international relations.

Luc Sindjoun
University of Yaoundé II
Yaoundé, Cameroon

See also Developing World and International Relations; Imperialism; Postcolonialism

Further Readings

Amin, S. (1974). *Accumulation on a world scale: A critique of underdevelopment.* New York: Monthly Review Press.

Amin, S. (1976). *Unequal development: An essay on the social transformation of peripheral capitalism.* New York: Monthly Review Press.

Ayoob, M. (1996). Subaltern realism: International relations theory meets the third world. In S. G. Neuman (Ed.), *International relations theory and the third world* (pp. 31–49). New York: St. Martin's Press.

Bhabha, H. K. (2007). *Les Lieux de la Culture* [Places of culture]. Paris: Payot.

Cardoso, F. H. (1973). Associated-dependent development: Theoretical and practical implications. In A. Stepan (Ed.), *Authoritarian Brazil: Origins, policies and future* (pp. 142–176). New Haven, CT: Yale University Press.

Chilcote, R. H. (1981). *Theories of comparative politics: The search for a paradigm.* Boulder, CO: Westview Press.

Frank, A. G. (1966, September). The development of underdevelopment. *Monthly Review, 18,* 17–31.

Lenin, V. I. (1967). *Selected works in three volumes.* Moscow: Progress.

Nkrumah, K. (1965). *Neo-colonialism: The last stage of imperialism.* London: Thomas Nelson & Sons.

Rodney, W. (1972). *How Europe underdeveloped Africa, London and Dar es Salaam.* Washington, DC: Howard University Press.

Roe, P. (1999). The intrastate security dilemma: Ethnic conflict as a tragedy. *Journal of Peace Research, 36*(2), 183–202.

Sindjoun, L. (2002). *Sociologie des relations internationales africaines* [Sociology of African international relations]. Paris: Karthala.

Development, Political

Political development gained recognition as a subfield of comparative politics in the 1960s. This entry describes its roots in modernization theory and dependency theory, examines broad themes in its subsequent evolution, and concludes with a look at the role of area studies in the discipline today.

Prior to the 1960s, comparative politics had tended to focus on the forms of government that prevailed in the advanced industrial world, that is, Europe and the former Soviet Union. But with the break up of Europe's empires following World War II and the entry of numerous postcolonial states into the global political system, an increasing number of scholars chose instead to focus on what became known as the developing world. Figure 1 shows the age composition of the states that became independent in different periods and captures the size of the new group of nations in the past 66 years: 120 out of 194, that is, almost two thirds.

The study of political development was foreshadowed by studies of nationalism. Initially, scholars treated the subject as a branch of contemporary history: The rise of nationalism in the postwar world seemed to resemble the rise of nationalism in Europe, be it in 19th-century Germany, Greece, or Italy, or in Central and Eastern Europe following the conference in Versailles. But many soon recognized that politics in the newly developed nations differed from that in the developed. They found it necessary to address new themes: economic growth, the rise of industry, and the marginalization of agriculture as well as the politics that accompanied these transformations.

While diverse intellectually, those making early contributions to the field tended to subscribe to what became known as "modernization" theory. Modernization theory constituted a claim that the breakdown of the previous agrarian and traditional order made people available for political mobilization. Increasing levels of urbanization, literacy, and income, its proponents argued, bore a close relationship with the rise of political participation. They posited sociological and social psychological mechanisms—the desire for community and the impact of anomie, for example—to link

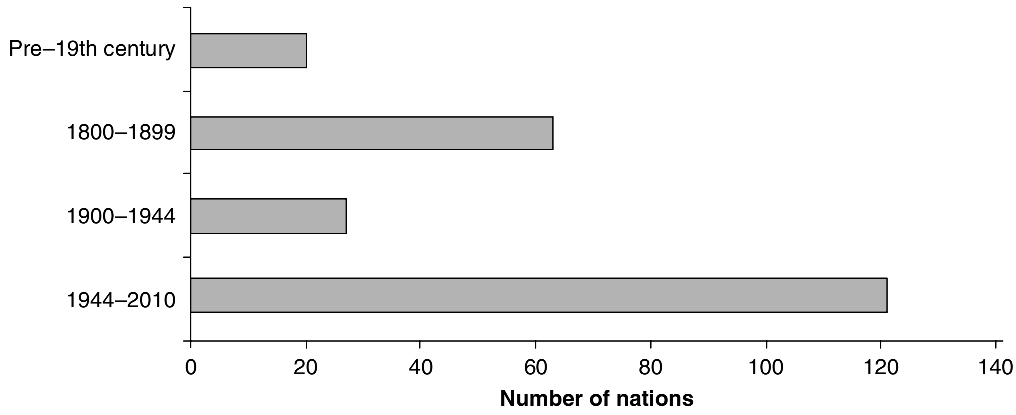


Figure 1 Date of Independence: Four Periods

economic and social change to the growth of mass movements. Drawing on the work of area specialists, several stressed the manner in which nationalist politicians harnessed the political energies of voluntary associations and in particular groups formed in town by recent migrants from the countryside. Rather than being highly organized and disciplined electoral machines, some nationalist parties, these specialists found, were but loose assemblages of these organizations—something that lent credibility to the reasoning of the modernization theorists.

The following is a key question faced by students of the developing nations: What kind of polities would emerge in the developing world? Given the bipolar structure of power in the postwar era, this question was equivalent to asking whether these nations would favor democracy or totalitarianism. When Seymour Martin Lipset reported a systematic relationship between the level of per capita income and the level of democracy in a cross-sectional regression of nation-states, the answer seemed clear. Given that higher levels of urbanization and literacy covary with higher levels of income and that—as Lipset found—higher levels of income correlate with democracy, then the developing world, as it modernized, was destined to be democratic.

This claim soon became the focal point of heated debate, however, one that brought to the fore an alternative form of political analysis. In the midst of the recession and debt crisis of the 1970s and 1980s, military governments overthrew democratic regimes in countries such as Brazil, Argentina, and Chile. As these nations numbered among the most prosperous in the developing world, when these regimes closed newspapers, banned political parties, and jailed or assassinated political opponents, they made a mockery of the claim that democracy covaried with development. Rather, their behavior lent credence to what became known as "dependency theory."

Modernization theory traced its intellectual roots to the writings of Max Weber, Émile Durkheim, and other sociologists (e.g., Ferdinand Tönnies) who studied the emergence of Europe as a modern society. The roots of dependency theory lay instead in the writings of Karl Marx, who focused on the economic forces that propelled Europe's 19th-century transformation. For important reasons, however, the dependency theorists did not draw on Marx himself but rather on his followers, in particular, Vladimir I. Lenin.

As Lenin recognized, the predictions of Marx had been falsified. In modern industry, wages of employees were rising; employers such as Ford found that by paying higher wages, they could sell more products. As a result, it was becoming increasingly difficult to write of the immiseration of the working class. Nor could observers find evidence of political conflict between the bourgeoisie and the proletariat. Thus, Lenin's acknowledgment that at the outbreak of World War I, in each major European nation, socialist parties had joined bourgeois parties in voting to finance their militaries. They were even prepared to kill workers if they wore the uniforms of another nation. The proletarian parties had acted as if they had more interests in common with their own nation's bourgeoisie than with another nation's working class.

Faced with these realities, Lenin turned to a key question: If the bourgeoisie were not exploiting workers in the advanced industrial countries, then from where did they draw their capital? The answer, he responded, was from abroad. Through multinational firms, international bankers, and global markets, the rich nations of Europe were able to extract resources from the poor nations of the periphery. Rather than taking place in the advanced industrial economies, the extraction of surplus took place at the global level.

For the dependency theorists, then, the fall of elected governments in the most prosperous nations in the developing world was not paradoxical. For the purpose of governments in the developing world, they believed, was not to serve the interests of their own people but rather the interests of investors from abroad. That the military governments repressed labor, destroyed trade unions, and jailed—or killed—socialist agitators was what was to be expected of any government occupying their position in the global system.

Lending further credence to the Leninist interpretation was the war in Vietnam. In the United States, workers—the so-called hard hats, in the argot of the time—supported the war; and in Vietnam, a peasant-based army fought the army of a capitalist nation. Political forces thus aligned in a manner redolent of Lenin's analyses. In addition, the war reminded scholars of Barrington Moore's claim that all modern revolutions were based on the peasantry rather than the proletariat—something that would follow from the reasoning of the dependency school.

Despite its initial appeal, dependency analysis fell victim to discordant facts. One was the rise of the industrialized economies in regions that formerly had been underdeveloped, most notably in Asia. Not only did these economies grow, and their nations escape the ranks of the underdeveloped, but also their growth was spurred by foreign trade—an instrument of immiseration, in most Marxian accounts. Also significant was the so-called third wave of democratization that culminated in the fall of the Soviet Union. With Russia reforming politically; Chile, Brazil, and Argentina restored to the ranks of democracies; and several of the newly industrialized nations of Asia choosing their rulers in contested elections, countries that had lain off Lipset's regression line snapped back on to it. Modernization theory appeared redeemed.

While Marxist political analysis may have lost its appeal, scholars continue to explore its central concerns: the impact of global trade, international finance, and the multinational structure of firms. And while democratic political forces may have displaced military and authoritarian regimes, many of the new democracies exhibit authoritarian features: legislatures with but token powers, limits on association and expression, and the maintenance in power of incumbent leaders or political parties. The result is a rapidly growing literature on "partial democracies." Contemporaneous with the spread of democracy in the developing world, moreover, was the growth of civil conflict, which leads to copious research into the causes of civil war and state failure in the developing world. As ethnicity and religion have played a major role in these conflicts, attention has refocused on the cultural roots of politics in the developing world. While each of these themes figured prominently in the early literature, scholars now approach them in new ways. Although the importance of cultural forces had long been recognized, for example, scholars now seek to find how, and under what conditions, they affect political behavior and collect data and design experiments to test their claims.

The Broad Themes

The decline of dependency theory left the field of development bereft of a "master narrative." While accepting many of the basic empirical findings of modernization theory, few scholars were inclined to subscribe to its tacit endorsement of the cultural, economic, or political triumph of the West. Their demurral became more pronounced following the late-century rise of a prosperous and powerful Asia. Rather than achieving a theoretical consensus, then, the field is marked by themes, which in turn lead to a redefinition of disciplinary boundaries.

Political Economy

Modernization and dependency theory focused on the political impact of economic change, linking (variously) democracy or authoritarianism to capital accumulation and economic growth. More recently, scholars have focused on the impact of politics on the performance of the economy. Some have looked at the role of interest groups and the way in which they shape policies toward trade, others on the regulation of markets and the manner in which they are transformed into political machines, and still others on the way in which the movement to democratic forms of government alter policy decisions, inducing or failing to induce them to employ public goods rather than private transfers in search of office. These lines of research have brought students of development into close contact with political scientists in the field of international political economy and economists studying politics.

Security Studies

Those studying development have also forged closer ties with those in the field of security. In recent decades, the locus of military security has shifted from the international to the domestic level, with far more deaths resulting from civil conflict than from wars. For years, students of international conflict have been able to apply geopolitical reasoning to "the state," viewing the latter as a unitary actor; the latter assumption is no longer tenable when studying conflicts at the subnational level. Just as students of development gain insights into the use of military power from their colleagues in security studies, those in security studies now find themselves benefiting from the insights of those who study the domestic politics of developing countries. Driven by the importance of this research theme, the subfields have come into greater contact.

History

While thus probing deeply into the politics of the developing world, scholars are simultaneously viewing it from a more detached vantage point. The recent rise of Asia reminds many that not so long ago, it was Asia that constituted the core while the West languished in the periphery. Inspired by this realization, students of development increasingly work in concert with students of economic history. Some return to the study of Western imperialism and the manner in which it, by toppling the nations of Asia, generated a reversal of economic fortunes on the one hand while laying the foundations for subsequent and enduring inequality on the other. Others look for factors, such as institutions, that can account for both differences in economic performance over time and across regions. Many look for lessons from history that might be applied to the developing world, focusing in particular on the origins of democracy and the role of warfare in rendering states "developmental."

Area Studies and the Discipline

A last major topic should be addressed: the relationship between those doing area studies and those who "do" social science, making extensive use of formal theory and empirical methods. Rather than treating the approaches as rivals, as has been the case in the past, scholars are now more inclined to treat them as complements. Drawn into political science because of their love of a region and their commitment to its development, scholars are increasingly inclined to combine a deep understanding of its politics and culture, derived from field research and the use of ethnographic methods, with the use of formal theory, statistical analysis, and experimental methods. The clash of cultures that once marked the field now appears muted; area studies are now firmly based within, rather than in opposition to, the discipline.

Robert H. Bates
Harvard University
Cambridge, Massachusetts, United States

See also Area Studies; Democratization; Dependency Theory; Developing World and International Relations; Development Administration; Economic Policy; Security and Defense Policy

Further Readings

Acemoglu, D., Johnson, S., & Robinson, J. (2001). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics, 117*(4), 1231–1294.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development. *American Economic Review, 91*, 1369–1401.

Coleman, J. S. (1958). *Nigeria: Background to nationalism*. Berkeley: University of California Press.
Coleman, J. S., & Rosberg, C. G. (1965). *Political parties and national integration in tropical Africa*. Berkeley: University of California Press.
Deutsch, K. W. (1961). Social mobilization and political development. *American Political Science Review, 55*(3), 493–510.
Emerson, R. (1960). *From empire to nation*. Boston: Beacon Press.
Lenin, V. I. (1970). *Imperialism: The highest stage of capitalism*. Beijing, China: Foreign Languages Press.
Lipset, S. M. (1960). *Political man*. Garden City, NY: Doubleday.
Moore, B. (1966). *Social origins of dictatorship and democracy*. Boston: Beacon Press.

Development Administration

Development administration refers to a form of public administration meant to be suitable for developing countries. The concept gained widespread recognition in the 1960s, coinciding with the emphasis at the time on state-led development. Development was something that needed to be centrally managed. The concept has since lost much of its initial value, and what many expected to be a new field of inquiry has never materialized. Instead, the concept has become subsumed under others such as development management and governance. This entry traces the history of and contending perspectives on the concept before concluding with a discussion of its demise and gradual inclusion into rivaling intellectual traditions.

History

Development administration emerged as a significant concept in the field of public administration in the late 1950s and early 1960s. Its promotion came from three distinct sources.

First, progressive civil servants who realized that the rigid system of administration that they associated with colonial rule would not work once these countries in Africa and Asia were independent constituted one constituency in favor of a more developmental form of administration. Several of these civil servants who themselves had personal experience of working in the colonial administration shifted to becoming international civil servants working on technical assistance contracts in these countries.

A second group was made up of philanthropic foundation officers with an interest in facilitating the development of new states in Africa and Asia. Among U.S. entities, the Ford and Rockefeller Foundations came to play an especially significant role. The former was very active in India and Pakistan as well as in Anglophone African countries. Support by these foundations was critical not only for improvement in administrative practices but also for experimenting with new theories of administration. The Ford Foundation–funded African Association for Public Administration and Management (AAPAM) played a particularly important role in fulfilling this dual objective in the 1960s and 1970s.

The third group consisted of academics with an interest in making a contribution to the cause of development by inventing new theories of administration that would be applicable in particular to the new states in the developing regions of the world. These scholars were largely from developed countries such as Australia, the United Kingdom, and the United States. This group eventually grew to include scholars from developing countries. Onkar P. Dwivedi from India, Jorge Nef from Brazil, and Dele Olowu from Nigeria are among the most significant contributors.

The enthusiasm with which the concept was met in the 1960s lasted into the 1970s but began to wane subsequently for two main reasons. The first was the shift in intellectual paradigm from a state-led to a market-inspired approach to development. Reforms in public administration became less urgent than similar actions in the economy. The second was the dearth of achievements in development administration. The theorists had largely failed to translate their ideas into practical action. What had been presented as a promising new academic field with practical value for developing countries had little to show for itself by the 1980s. What happened thereafter was an attempt to redefine reform in broader terms than had been the case in the previous 2 decades.

Contending Perspectives

There has never been a single and coherent theory guiding the field of development administration. It failed to develop beyond a preparadigmatic stage. There were four main theoretical perspectives that competed for domination of the field, although none ever reached that point.

The first theoretical effort by Victor Thompson drew its inspiration from an earlier study of enterprise management in the United Kingdom. The authors of this study made a distinction between two forms of management—mechanistic and organic. The former would coincide with a classical bureaucratic form in which authority follows hierarchical lines, roles are clearly defined and rules strictly enforced. The latter relies on a more flexible form of organization in which relations among employees are more collegial, and rules can be bent in the interest of achieving a common overarching goal. Thompson borrowed these ideas and developed his own theory of organization that applied to how administration should be structured to be conducive to development. His assumption was that since organic management was especially suitable in conditions of change, for example, in technology or market opportunities, this could be applied also to administration of development since the latter implies change. Thompson's theory focused on how behavioral change could be induced by senior administrators. It assumed that bureaucratic organizations could be reformed from within. His theory, therefore, was attractive to practitioners interested in making public administration in the new states more developmental.

A second approach that is associated especially with Fred Riggs started from a systems point of view and was less precise in terms of what could or should be done. His "prismatic theory of society" was inspired by modernization theory, according to which societies move from being "traditional" to becoming "modern." Riggs's contribution was that in the field of administration, states in Asia, Africa, or Latin America were caught somewhere in between, with influences coming both from the past and the present, hence creating a syncretic combination that was different but not necessarily developmental. His work attracted interest from other scholars as long as modernization theory had its influence until the early 1970s but lost it immediately thereafter.

Whereas Thompson had concentrated on what could be done with development administration theory inside organizations, Riggs argued that administrative organizations reflected the values intrinsic in society. Theories of development administration could not be applied without attention to its "ecology" or societal environment. These contending perspectives dominated the debate in the 1970s. David Leonard made perhaps the most original and empirically based contribution to this debate with his pathbreaking study of the agricultural extension services in Kenya. He reached two important conclusions that helped set the debate into the 1980s. The first was that rejecting a mechanistic form of organization in favor of an organic one had its own costs because there are certain basic functions in bureaucracies, such as personnel management and accounting, that cannot be overlooked if administration is going to be developmental. Preferring the concept of "development bureaucracy," Leonard suggested that there is not necessarily a tension between mechanistic and organic forms of organization. The second contribution was that although societal factors matter, organizations can make a difference if properly administered. Government ministries are not "hopeless cases" but rather potential change agents. Leonard, therefore, kept the field of development administration alive by merging previously contending perspectives and arguing that development is possible through more careful and systematic attention to bureaucratic organizations.

A fourth perspective on development administration was provided by George Gant, who made a distinction between what may be called a more down-to-earth administration in specific-line ministries, on the one hand, and a comprehensive development administration concentrated in a superior branch of government, for example, a President's Office or a Planning Commission, on the other. Gant was skeptical that a developmental form of administration could emerge in day-to-day bureaucratic settings and advocated, therefore, the creation of a "super" administration in charge of development, an idea that he appears to have borrowed from the literature and practice of development planning.

This account of significant contributions to the field would not be complete without mentioning two other names. The first is Milton Esman, who through various contributions followed the evolution of the field from its heyday to its demise. Together with colleagues at Cornell University, he was also responsible for applying organizational theory to empirical reality in South Asia. The other is Bernard Schaffer, who wrote a seminal piece in 1969 in which he was the first to criticize the ambitions of the field for being naïve about the possibility of making bureaucratic forms of administration more developmental. Drawing especially on his studies and experience in India, Schaffer argued somewhat along the same lines as Riggs that where norms in society are different from those required in a functioning administrative organization, development administration is unlikely to emerge.

Demise

Development administration never took off as a field of its own. It became a victim of its own high aspirations and the difficulties of realizing them. The shift in the 1980s away from state-led development exacerbated this difficulty. What has happened since may be captured in two distinct trends. One has been the incorporation of development administration into the field of public administration—a return to where it all came from in the first place. This manifests itself in the belief that for developing countries to become successful, they require a public administration similar to that found in developed countries. The multiple public sector reform programs that have been launched in the 1990s and 2000s with generous funding from the donor community are a confirmation of this trend. In this perspective, development administration has been reduced to simply being administration in developing countries. It has no longer a theory of its own.

The second trend has been the move from development administration to governance. This metamorphosis has subordinated issues of public administration to a broader set of concerns, both economic and political. Thus, rather than treating issues of administrative reform as a separate concern, these issues are interpreted in a broader economic and political reform perspective. Costs of administration have become a central issue thanks to its encounter with economics. Likewise, administration and politics are no longer two distinct fields but brought together in the name of good governance to make both activities more relevant and appealing to the public, an objective that is now applied by both scholars and development practitioners to developed and developing countries alike.

Goran Hyden
University of Florida
Gainesville, Florida, United States

See also Administration; Administration Theory; Development, Political; Economic Policy; Governance, Administration Policies

Further Readings

Gant, G. F. (1979). *Development administration: Concepts, goals and methods*. Madison: University of Wisconsin Press.

Leonard, D. K. (1977). *Reaching the peasant farmer*. Chicago: University of Chicago Press.

Riggs, F. (1964). *Public administration in developing countries: The theory of prismatic society*. Boston: Houghton Mifflin.

Schaffer, B. (1969). The deadlock of development administration. In C. T. Leys (Ed.), *Politics and change in developing countries* (pp. 177–212). Cambridge, UK: Cambridge University Press.

Thompson, V. (1964). Administrative objectives for development administration. *Administrative Science Quarterly, 9,* 91–108.

Diaspora

The word *diaspora* comes from the ancient Greek *dia speiro*, meaning to sow over. It refers to populations that originated from the same place but have now scattered to different locations. The concept of diaspora has long been used to refer to the Greeks in the Hellenic world and to the Jews after the fall of Jerusalem, and beginning in the 1950s and 1960s, scholars began to use it with reference to the African diaspora. However, it has come to be used more widely during the past 2 decades. This entry first puts the concept of diaspora in historical perspective and discusses the various typologies that researchers have developed in analyzing it. The entry then focuses on the political implications of both the phenomenon and its interpretations, and finally, it considers what is at stake in the transformation of the nation-state system that arises out of the diaspora experience.

Evolution of the Concept of the Diaspora

The concept of diaspora did not figure prominently in the social sciences until the late 1960s, and the use of the plural is recent. It used to refer primarily to the Jewish experience, in Greek versions, particularly through the expulsion of the people and destruction of the Jerusalem temple under the Babylonian empire. The Jewish population's experience of a dispersal made necessary by its loss of territory shapes a tragic vision of the diaspora, long shared by many analysts. Nonetheless, since ancient times, the concept has been used in a positive though much less influential way to refer to the Greek colonization between the 6th and 4th centuries BCE in all the Mediterranean lands, from the shores of present Turkey and Crimea to the Strait of Gibraltar, spreading civilization in these territories through many Hellenic cities.

Both experiences—rooted in the Western tradition—have constituted stereotypes of diasporas, though other noticeable cases from the East have developed for the medieval and modern times. For instance, the expansion of China has been perceived as an acceptable phenomenon, described in an ancient Chinese poem: "Wherever the ocean waves touch, there is an overseas Chinese"—thus naturalizing diasporic initiatives along trade routes. India's expansion, especially throughout the Indian Ocean, has also provided an example of the settlement of a population beyond its own boundaries. Since the 19th century, the increase in the populations of coolies to work in agricultural or industrial plants worldwide has drawn particular attention.

In fact, when scholars attempted to classify different experiences, from the late 1980s onward, in order to be able to identify diasporic processes beyond the unique and restrictive Jewish case, the Greek, Chinese Indian, and African ones served to establish a typology, along with the former. As Robin Cohen (1997) describes, diasporas were classified as victim, imperial/colonial, trade, or labor diasporas, according to the main motives that generated original migration—namely, expulsion, expansion, commerce, or work. Other examples, fitting into such types, demonstrated the transhistorical relevance of such a classification. For instance, the Armenian exile in Europe, British thalassocratic empire, Lebanese trading posts in Africa and Latin America, and Moroccan cheap labor settlements in Western Europe.

Other typologies, departing from functionalist views, emphasize historical or political factors. Thus, according to Gabriel Sheffer (2003), they mention traditional/historical (Jewish, Greek, Phoenician), dormant (American/U.S.), incipient (Russian, post-Soviet), stateless (Palestinian, Roma/gypsy), and state-linked (Mexico, China) diasporas. Such classifications assume that diasporas are unstable social bodies whose morphology can change over time, passing from one category to another. This evolutionary and flexible approach shows that, if diasporas have indeed existed since ancient times, some have formed, others have disappeared, but even more have been recently created.

Most of the scholars converge on the fact that massive population moves since the middle of the 19th century have generated multiple diasporas that became especially visible in the late 20th century. As a world map of the impact of migrations would show, durable expatriate communities have been established around the globe. The question that arises is, Do these settlements share enough common points to be all denominated diasporas? The issue of criteria divides the academic community. Some scholars insist on permanence (of the settlements), others on internal community organization, on the contrast with or exclusion from host societies, on social composition of expatriate communities, on the modes of communication between these, on the number of dispersed units, or on the will to come back to the place of origin, and so on. According to the evidence they collected, they argue that one or several of these criteria will prevail over the others and define a true diaspora.

Political Significance

The minimal feature of diasporas is, in fact, the dispersion from a common origin. This origin may be partly mythical or purely symbolic, as in the case of the Black/African diaspora, whose collective identity resides more in a sociocultural experience of suffering (based on a history of exploitation) than in an actual relationship (Stéphane Dufoix, 2003). However, most of the diasporas have kept relationship with the place of origin and among the scattered groups themselves. As Georges Prévélakis (1996) notes, the network structure that links the different parts is therefore a general feature. According to Sheffer (2003) and William Berthomière and Christine Chivallon (2006), because the origins of recent diasporas are existing or potential nation-states, some authors qualify these as ethno-national diasporas to explicitly distinguish them from transnational networks in general that have developed in the context of globalization. In that sense, diasporas appear as transborder extensions of nation-states and deterritorialized countries' populations, in an increasingly normal pattern.

Today's estimate is that almost 1 out of 10 human beings is living in a diasporic situation (between 400 and 600 million). The number of individuals with dual citizenship has exploded in a short period of time. In Latin America alone, from 4 countries allowing it in the 1980s, the number had reached 10 by early 2000 (Brazil, Colombia, Costa Rica, the Dominican Republic, Ecuador, El Salvador, Mexico, Panama, Peru, and Uruguay; see Michael Jones-Correa, 2001).

Surveys show that dozens of countries today have set up organizations, institutions, procedures, and devices of all sorts to reach and capitalize on their expatriates. Financial remittances of migrants (not only first generation) have gone up to several hundred billion dollars per year and are increasingly channeled for productive collective projects, not just for individual consumption purposes. Social remittances such as technology transfers, information or knowledge exchanges, organizational/management norms, or democratic values transmission are also mentioned as positive inputs for development in home countries. Migrants' and expatriates' associations are burgeoning in many host countries.

The emerging interest of diaspora populations in their countries of origin leads host countries to worry about whether diaspora populations may manipulate or exploit them. These host countries fear a fifth column operating against national interest or suspicious ethnic networks involved in delinquent or terrorist activities. However, the host countries' approach is generally favorable to diasporas and supportive of their organizations. They consider that it is easier to control migrant groups through these organizations than to deal with amorphous and heterogeneous populations without proper representation. At the same time, cooperation through diaspora groups creates all sorts of opportunities abroad for the receiving nations.

Cooperation, however, cannot be taken for granted. In a number of cases, diasporas come from origin countries where their members are not welcome and where free circulation is limited, making cooperation impossible. On the other side, xenophobia and a reluctance to accept foreign people have not disappeared and can spread in crisis situations. Diasporic conditions in a nation-state system remain submitted to territorial legislations and decision making, which may limit their initiatives.

Future Perspectives

The diaspora may sometimes appear as a step toward a postnational system of governance, as a transition to a cosmopolitan regime. The fact that nation-states are no longer contiguous, but rather expand in networks in which national boundaries overlap, does change the landscape of international relations. So far, it seems that this generally does not bring problems. The fear that double citizenship, for instance, would create a conflict of loyalty between the local/host government and the country of origin has not materialized. The reason is simply that the probability of irremediable antagonism between both poles of migration is extremely low. When it happened in the past, in exceptional cases, expatriates tended to identify with the country where they live, work, and raise children. Generally speaking, today, individuals and groups from the diaspora combine identities, feeling that they belong to both home and host countries and that they can mix both easily in their daily life in a nonexclusive and productive manner.

Today, many people claim to be living in a diaspora, to be part of a minority, or to have ancestors from a different ethnic group. Contrary to former pessimistic views, they attach a positive value to this as if it added a premium of identity, rather than a negative stigmatization. The present-day mobility of individuals no longer means an experience of being uprooted and immersed in a foreign world; rather, it is seen as the ability to keep in touch with relatives and contacts abroad/at home and to remain connected to cultural, cognitive, and symbolic values of remote places. Information and communication technologies have obviously facilitated this new proximity, but host countries' evolution from a homogeneous conception of citizenship toward more pluralistic, multiethnic approaches is also crucial. Today, more than in the past, political, and socioeconomic integration may be dissociated from cultural and relational assimilation.

Some authors consider this as a postmodern condition in which individuals have multiple and fragmented identities that the diaspora would be reflective of. There are indeed strong examples of hybridization through the diaspora—for instance, expatriate engineers and information scientists joining associations of Asian natives in North America to promote their mutual interests and careers through social networking. Ethnic background thus becomes a means for achieving economic goals in a foreign environment, with an individual's multifaceted affiliation serving as a professional advantage. In that sense, the very term *ethno-national diasporas* may be restrictive, today, as it reduces the expatriates groupings to a single dimension. A social-constructivist approach, in which identity building results from local interactions and interests, articulated with larger determinants through network-operated negotiations, is appropriate. Identities are thus no longer petrified in macrosocial categories, but neither have they vanished in an atomistic world.

In such a situation, the nation-state still remains sovereign. According to Sheffer (2003), until now, there has been no case in which the diaspora has imposed its own decisions on the country of origin, even though there have been many effective interventions of expatriates (e.g., the struggles for democratization in South Africa and Eastern Europe in the 1990s). At the opposite end, direct and decisive involvement in nation-states has been seen in the actions of the former Soviet Union toward incipient Russian diasporas. This reminds us that new actors and relations, such as diasporas, may be translated and interpreted in traditional politics, where irredentism might be perceived again as a source of problems, in a purely realistic approach to international relations.

Jean-Baptiste Meyer
Institut de Recherche pour le Développement
Montpellier, France

See also Constructivism in International Relations; Globalization; Identity, Social and Political; Migration; Networks; Postmodernism in International Relations

Further Readings

Berthomière W., & Chivallon, C. (Eds.). (2006). *Les diasporas dans le monde contemporain* [Diasporas in a contemporary world]. Paris: Karthala/MSHA.

Cohen, R. (1997). *Global diasporas: An introduction*. London: UCL Press.

Dufoix, S. (2003). *Les diasporas* [The diasporas]. Paris: Presses Universitaires de France.

Jones-Correa, M. (2001). Under two flags: Dual nationality in Latin America and its consequences for naturalization in the United States. *International Migration Review, 35*(4), 997–1029.

Prévélakis, G. (Ed.). (1996). *The networks of diasporas*. Nicosia, Cyprus: Cyprus Research Centre (KYKEM).

Sheffer, G. (2003). *Diaspora politics at home and abroad*. Cambridge, UK: Cambridge University Press.

Dictatorship

A dictatorship is a form of government in which a person or a group has absolute power, unlimited by constitution or other laws and not based on traditional legitimacy. The meaning of the term, and the content of the phenomenon itself, has changed over time. In political science, the study of dictatorship constitutes an important subject. Understanding it is vital for comparative politics and for the study of new democracies that emerged from the ashes of former dictatorial regimes. This entry first describes the history of dictatorships and reviews the major typologies used to classify them. It then discusses the causes of dictatorships and the various ways in which they have been ended.

History

The term *dictator* was used for the first time in ancient Rome. In times of calamities, consuls, on the recommendation of the Senate confirmed by the popular assembly (*comitia curiata*), appointed a dictator (with the title of *magister populi*) from among former consuls. Such dictators had unlimited power (*imperium summum*) for a limited period of 6 months. A dictator was first appointed in 501 BCE. In the 1st century BCE, the institution changed its character; under the dictators Lucius Cornelius Sulla (82–79 BCE) and Gaius Julius Caesar (49–44 BCE), dictatorship became the permanent rule of powerful individuals. After the abolishment of the Republic in 33 BCE, the Roman Empire became an early version of military dictatorship with the Praetorian Guard holding effective power to appoint and overthrow the emperors. Late Roman praetorianism became a prototype for the military dictatorship in modern times. The collapse of the Roman Empire resulted in the disappearance of the dictatorial system for more than 12 centuries.

The concept of dictatorship appeared in political thought in the 16th century. Niccolò Machiavelli, in the 34th chapter of *Discourses*, criticized the Roman dictatorship for having led to tyranny, but he advocated the idea of a dictator ruling by the will of the people to bring unity and freedom to Italy. Jean Bodin used the term *dictatorship* to define sovereignty, and Jean-Jacques Rousseau defined dictatorship as a system suspending the legitimate government.

Early-modern dictatorships were the result of revolutions, which abolished the monarchy but were unable to establish stable democratic governments. The rule of Oliver Cromwell in England (1649–1658) and Napoleon Bonaparte in France (1799–1814) are the best known examples of successful military commanders becoming absolute rulers due to the weakness of the democratic institutions in their respective countries and the support of the armed forces under their command. Following the national revolutions in Latin America in the early 19th century, dictatorial systems headed by military commanders (caudillos) were established, and in many countries of the region, they became the dominant pattern of government until the second half of the 20th century.

Europe faced the strongest wave of dictatorships in the first half of the 20th century. The Russian Revolution of 1917 resulted in the establishment of the first dictatorship of the Communist Party. Its ideological justification was based on the Marxist concept of the "dictatorship of the proletariat," coined by Karl Marx but redefined by Vladimir Lenin. According to the Leninist doctrine, during the long process of constructing the communist society, power should be concentrated in the hands of the Communist Party, defined as the avant-garde of the working class.

The collapse of the Ottoman Empire resulted in the establishment of a secular one-party regime in Turkey under the leadership of Mustafa Kemal Pasha.

During the interwar period (1918–1939) in all newly independent states of Central Europe with the exception of Czechoslovakia, democratic governments were overthrown and dictatorships, mostly military, were established. In 1922, the fascist party of Benito Mussolini seized power in Italy; in 1926, a dictatorial regime was established in Portugal; and in 1933, the National Socialist Workers Party of Germany (Nationalsozialistische Deutsche Arbeiterpartei, or NSDAP) headed by Adolf Hitler came to power in Germany. After winning the Spanish Civil War (1936–1939), the military under Francesco Franco established a dictatorial system based on a coalition of the military, the Roman Catholic Church, and the fascist party Falange. On the eve of World War II, the majority of European states were ruled by dictatorial regimes. The Japanese military established their de facto power behind the facade of the absolute rule of the emperor.

World War II resulted in the defeat of dictatorial regimes in Germany, Italy, and Japan, opening the road to the democratization of those three countries. In East-Central Europe and North Korea, however, the hegemony of the Soviet Union resulted in the establishment of dictatorial rules of the communist parties. In later years, communist dictatorships have been established through revolutions and wars for independence in China (1949), Cuba (1959), Vietnam (1954 in the North and 1975 in the whole country), and Cambodia (1975). In Western Europe prewar dictatorships survived only in Portugal (until 1974) and Spain (until 1975). In Greece, the military established dictatorship in 1967, only to abdicate its power in 1974 after the abortive attempt to incorporate Cyprus.

The process of decolonization produced several dictatorial regimes in the new states of Africa where the democratic governments established by the departing colonial powers proved to be weak and inefficient. The same process took place in some newly independent states of Asia (Pakistan, Burma, and South Korea). In the Arab countries of North Africa and the Middle East, nationalistic dictators came to power following the pattern of the Egyptian revolution of 1952. As the result of these processes, in the early 1970s, dictatorships of various types ruled over more than half of the world population. Of the 122 states with more than 1 million inhabitants, only 30 (24.6%) were democracies in the 1970s (data from Freedom House definition with regard to civil and political rights). Democratic governments survived in Western Europe, North America, and Australia and in a few other states.

The retreat of dictatorship began during the 20 years of the third wave of democratization (1974–1995). Between 1974 and 1975, Portugal, Greece, and Spain abolished dictatorships. In the 1980s, democracy was reestablished in most of the Latin American states, and in 1989, the process of democratization started in the communist states of East-Central Europe, followed by the abolishment of the communist rule and the eventual dissolution of the Soviet Union. In the late 20th and early 21st centuries, there were also democratic changes in some countries of Africa and Asia. The net result of the third wave of democratization was the reduction in the number of dictatorial regimes. In 2008, according to the Economic Intelligence Unit, dictatorial regimes survived only in 51 (30.5%) states with 34.9% of the world population. The largest and most powerful contemporary dictatorship is the rule of the People's Republic of China by the Communist Party of China. Communist dictatorships survived also in North Korea, Vietnam, and Cuba. In some post-Soviet republics in Central Asia and the Caucasus (Azerbaijan, Kazakhstan, Tajikistan, Turkmenistan, and Uzbekistan), authoritarian rule substituted for the former communist dictatorship. Dictatorships survived also in several countries of Africa and Asia.

Typologies

Modern dictatorships differ in their political form, social content, and ideological orientation.

Political Form

As far as the political form of dictatorship is concerned, the most important distinction introduced in the early 1950s, mostly by scholars such as Hannah Arendt, Carl J. Friedrich, Karl Wittfogel, and Zbigniew K. Brzezinski, was between totalitarianism and other forms of dictatorship. Later, Juan J. Linz and Leonardo Morlino developed this typology by comparing *totalitarian* and *authoritarian* types of dictatorship. According to them, totalitarian dictatorship differs from other forms of dictatorial rule by the existence of the "totalitarian syndrome" composed of

1. an ideology professing the total transformation of social relations,
2. the rule of a single party opposition to which is illegal,
3. state monopoly of the means of violence and of the media,
4. state control of the economy, and
5. police terror.

Two of the most important forms of totalitarian dictatorship have been fascism and Nazism on one hand and communism on the other. Authoritarian regimes make up a more diversified category, with military dictatorships being the most common but not the only form of such a system. They also tend to be less stable than totalitarian regimes.

A different typology related to the political form of dictatorship focuses on who holds power in a dictatorial regime. From this perspective, dictatorships can be divided into three main categories: *personalistic*, *military*, and *party* dictatorships.

Personalistic dictatorship is characterized by the concentration of power in the hands of the supreme leader who dominates the institutions. Totalitarian dictatorship took this form under powerful leaders such as Hitler, Stalin, Mao Zedong, and Fidel Castro, who dominated the political institutions. Military regimes based on the domination of armed forces were the most common form of dictatorship in the 19th century, particularly in Latin America. Military dictatorships in the early 21st century include those in Myanmar (Burma), Fiji, Gambia, Ivory Coast, Democratic Republic of Congo, Libya, Mauritania, and Syria, as well as authoritarian regimes established by military coups (in Egypt, Tunisia). Party dictatorships have had two main forms: totalitarian (fascist, Nazi, and communist) and non-totalitarian (Turkey 1923–1945, Mexico 1929–1977). Some communist dictatorships (Yugoslavia after 1949, Poland after 1956) lost their totalitarian character, remaining authoritarian versions of party dictatorship. There have been several dictatorships based on a mixture of military, party, and bureaucratic characteristics.

The German political writer Carl Schmitt introduced a third typology that distinguished between two main types of dictatorship based on the political form of dictatorship: *commissarial* and *sovereign*. The first is established to restore order so that the existing constitution can function normally, while the latter abrogates the existing constitution to establish a new political order.

Social Content

A typology based on the relation between the dictatorship and the social structure distinguishes between three types of dictatorial rule: *conservative*, *modernizing*, and *counterrevolutionary*. The conservative dictatorship serves the interests of the dominant social classes and defends the existing social order. A peculiar version of the conservative dictatorship, subordinated to the Islamic clergy and devoted to the preservation of religious values, was established in Iran in 1979 and in Afghanistan in 1992. The modernizing dictatorships attempt to reform the society from above, sometimes with the support of the underprivileged strata of the society. The counterrevolutionary dictatorship came to power as the result of a successful use of force in defense of the interests of the privileged strata endangered by the radical reforms of a modernizing government. The great majority of the military dictatorships belong to the conservative category, but some have been modernizing regimes (e.g., Turkey under Mustafa Kemal Pasha, 1923–1938; Egypt under Abdel Nasser, 1952–1970; and Argentina under Juan Peron, 1945–1955). Spain under Francesco Franco (1936–1975) and Chile under Augusto Pinochet (1973–1989) are classic cases of counterrevolutionary dictatorships. They were established by military coups to undo the socialist reforms carried out by democratically elected governments. With the passing of time, the social role of the dictatorship often changes, with the regime becoming more the guardian of the vested interests of the ruling elite rather than a champion of social change. Such change took place in the communist regimes.

Ideological Orientation

Typology based on the relation between the regime and ideology distinguishes between *ideological* and *pragmatic* dictatorships. In addition to totalitarian dictatorships, which have been ideological by definition, there have been some other ideological dictatorships, but most dictatorial regimes are based rather on vaguely defined ideas, such as the national interest or the political legacy of the founder of the regime (e.g., Kemalism in Turkey, Peronism in Argentina), than on more articulated ideologies. In addition to communism, fascism, and Nazism, which were the ideologies of the totalitarian regimes, religious ideologies and nationalism have been the most frequent types of justification for modern dictatorships. Spanish and Portuguese dictatorships relied on Roman Catholicism as the ideological rationale of the regime, while Islam has been the ideological base of dictatorships in several countries in the Muslim world, particularly in Iran since 1979 and in Afghanistan from 1992 to 2001.

Causes and Mechanisms of Dictatorship

Explanations of the emergence and survival of dictatorships refer to (a) political and societal conditions that make it possible to abolish the existing system and to establish a dictatorship, (b) motivations of the principal actors, and (c) techniques employed by the architects of dictatorships.

Political and Social Causes

The main political cause of dictatorships is the weakness of the former regime. Such a political situation results from a variety of causes, such as (a) the weakness of traditional monarchy, no longer commanding the loyalty of its subjects; (b) weakness, ineffectiveness, and corruption of the democratic governments combined with the lack of democratic mechanisms to change such a situation; (c) the lack or weakness of democratic political culture, both at the elite and at the mass level; (d) humiliation and frustration caused by the defeat in a war; and (e) imposition of a dictatorial regime by the hegemonic power (as was the case in most of the Communist dictatorships in East-Central Europe after World War II). Social causes of dictatorship, often combined with the political ones, include the following: (a) sharp social conflicts endangering the existing social and economic system, (b) economic crisis resulting in a drastic deterioration of the standard of life of the large part of the society, and (c) the emergence of a large number of people (mostly young males) who have lost, or never gained, a stable place in the social structure and are likely to become foot soldiers of a movement to establish dictatorship.

Motivations

Motivations to establish a dictatorship include ideological or religious beliefs, frustration with the existing political and/or economic situation and the feeling that extraordinary means are necessary to save the state, group interests (particularly the corporate interests of the military), and the personal ambitions of a strong leader.

Techniques

In most cases, techniques of the establishment of a dictatorial regime are based on the use of force, particularly military force. Such a coup takes various forms: from the threat to use force (pronunciamento in Latin American terminology) through the forceful removal of the existing government to a prolonged civil war. In the first two cases, cohesion of the forces supporting the coup is crucial for its success. In the case of a civil war, its result depends on a number of military and nonmilitary factors, particularly on the ability to win the support of a considerable part of the population.

Newly established dictatorships often seek a degree of legitimacy. Some rely on plebiscites in which citizens ratify the results of the coup. Napoleon Bonaparte first introduced the plebiscitary form of legitimating a dictatorship. Some regimes use a version of noncompetitive (plebiscitary) election to gain a degree of legitimization. Mass propaganda campaigns, sometimes with direct participation of the dictator, serve also to legitimize the regime. The effectiveness of such campaigns varies. Some dictatorships have been able to obtain a strong, albeit nondemocratic, mandate, while others are forced to rely mostly on naked coercion.

Dictatorships differ in the degree to which they use coercion. The common characteristic of dictatorships is that all of them use coercion to suppress opposition and/or to achieve other goals. However, the degree and nature of coercion vary.

In their formative stage, all dictatorial regimes tend to use repression against real or alleged enemies, but the intensity of such repression varies. Thousands of republicans were killed after the Spanish Civil War, and more than 3,000 supporters of the overthrown government were exterminated by the Chilean military after the coup of 1973. There were, however, relatively benign dictatorships, which used coercion rarely and tolerated independent political activities, which were severely punished in other dictatorships (Poland after the military coup of 1926). In most cases, the prolonged existence of the dictatorships leads to the reduction of the intensity of coercion and, in many cases, to the gradual liberalization of the regime. However, even in a relatively benign dictatorship, the state of human rights departs significantly from the democratic standards.

Extrication From Dictatorship

The great majority of modern dictatorships have existed for relatively short periods of time, usually less than 2 decades. Authoritarian dictatorships in Portugal (1926–1974) and Spain (1936–1975) were exceptionally durable. In Egypt after the 1952 military coup, the problem of succession in an authoritarian dictatorship was twice resolved by the orderly access to power of former vice presidents, Anwar Sadat in 1970 and Hosni Mubarak in 1981, following the death of their predecessors; however, Mubarak's regime was overthrown in 2011. Communist dictatorships lasted much longer than most of the noncommunist dictatorial regimes. The Soviet system survived 74 years and the communist regimes in East-Central Europe for more than 40 years. The still existing communist dictatorships in China, Cuba, North Korea, and Vietnam belong to the oldest dictatorial regimes in the contemporary world. Even they, however, are much younger than old democracies, which function successfully for many generations, in some cases for more than 2 centuries (the United Kingdom

[UK], the United States) and in many for more than a century.

The relative instability of dictatorships results from several structural weaknesses. Dictatorships usually have problems in establishing secure mechanisms of succession. This is particularly true in personalistic dictatorships since as long as the supreme leader is alive nobody is allowed to become the official successor. Selection of competent functionaries is difficult because political loyalty is considered more important than professionalism. Dictatorships also have problems with handling popular dissatisfaction. Since there are few possibilities of legal expression of protest, dissatisfaction leads to alienation from the regime and to spontaneous outbursts. When the level of such dissatisfaction is too high, naked force is not enough to secure the survival of the regime. After World War II, the international environment made it more difficult to maintain dictatorial systems of government in those parts of the world where the influence of the democratic powers were the strongest.

There have been several ways of extrication from dictatorships. Not all of them, however, were fully successful in the sense of establishing stable democracies. In several cases, departure from a dictatorial system led to establishment of another dictatorship. In some others, the result was unstable democracy or even anarchy and the collapse of state power.

The main patterns of extrication from dictatorship are (a) defeat in war and foreign occupation, (b) revolution or a coup, (c) democratization of the regime from above with or without negotiations with the opposition, and (d) capitulation caused by a major setback for the existing regime.

Military Defeat and Foreign Occupation

Some dictatorships have been abolished as a consequence of the war lost. Napoleon Bonaparte lost power after his defeat in the anti-Napoleonic Wars of Liberation (1813–1814). Several European dictatorships, including the German and Italian ones, as well as the military rule in Japan, were destroyed as a consequence of World War II. The dictatorship of the Red Khmers in Cambodia was removed from power by the invading army of Vietnam in 1978. In Afghanistan in 2001 and in Iraq in 2003, dictatorial regimes were removed by the military intervention of the United States and its allies. The removal of a dictatorship by foreign forces opens the door for the construction of a democratic system under temporary foreign protectorate, in some cases lasting several years. Not always, however, was the process as successful as it had been in Germany, Italy, and Japan.

Revolution or Coup

Sometimes dictatorships are overthrown by a revolution or by a military coup. Such cases were few, due to the effective use of force by the dictatorial regimes and the relative weakness of their opponents. The instances of successful revolutions launched against the dictatorship regimes were Cuba in 1959 and Portugal in 1974. Originally, a popular revolt in Hungary in 1956 managed almost to destroy the communist dictatorship, but soon it was crushed by the military intervention of the Soviet Union. Only in one case—Romania in 1989—a communist dictatorship has been overthrown by popular reform, supported by the armed forces. Revolts and coups often result in one dictatorship being substituted by another. In several African states, dictators were overthrown by armed rebels who then established their own dictatorial rule. In some Latin American states (e.g., Argentina in 1955), the military removed dictators only to establish their own rule.

Democratization

In several cases, dictatorships have been abolished through a process of democratization initiated by the reform-oriented wing of the existing regime, with or without cooperation with the moderate wing of the opposition. In Turkey, the ruling party initiated the democratization in 1945 by permitting the formation of opposition and by introducing competitive elections. Brazilian democratization from above (*abertura*, a Portuguese term that means “opening,” used to describe Brazil’s transition from military to elected government) began in the second half of the 1970s and resulted in full return to democracy in about 10 years. A similar process took place in Mexico where the new law of 1977 opened possibilities for electoral competition between the ruling party and the opposition and in the Republic of Korea in the last years of the 20th century. In Spain, after the death of General Franco in 1975, King Juan Carlos opened the process of negotiations with the (then illegal) opposition, which resulted in restoration of democracy in few years. In Poland in 1989 and in Hungary in 1990, negotiations between the ruling communist parties and the opposition led to competitive elections and to restoration of democracy. Similar processes of democratization from above took place in some republics of former Yugoslavia (Slovenia, Croatia, and Macedonia) as well as in Bulgaria, Albania, and Mongolia. In the Soviet Union, democratic reforms were initiated by the last leader of the Communist Party, Mikhail Gorbachev, in the second half of the 1980s and resulted in the collapse of the communist system in 1991.

Capitulation

Abdication of a dictatorial regime takes place when a serious external fiasco or a peaceful mass protest makes the status quo untenable. In Greece, the military government abdicated its power after it had provoked an international crisis over a failed attempt to incorporate Cyprus in 1974, and in Argentina, the military rulers gave up their power after the humiliating defeat in the short war with the UK over the Falkland Islands in 1983. The authoritarian regime in Philippines collapsed in 1986 as the result of mass protests, supported by the Roman Catholic Church and the armed forces. Communist governments in Czechoslovakia and in the German Democratic Republic capitulated in late 1989 when they had been confronted with mass protests and were unable to suppress the protest due to the lack of Soviet support.

Jerzy J. Wiatr
European School of Law and Administration
Warsaw, Poland

See also Authoritarian Regimes; Democracy, Types of; Democratization; Hybrid Regimes; Totalitarianism; Transition

Further Readings

Arendt, H. (1951). *The origins of totalitarianism.* New York: Harcourt & Brace.

Finer, S. E. (1962). *The man on horseback.* London: Pall Mall Press.

Friedrich, C. J., & Brzezinski, Z. K. (1956). *Totalitarian dictatorship and autocracy.* Cambridge, MA: Harvard University Press.

Huntington, S. P. (1991). *The third wave: Democratization in the late twentieth century.* Norman: University of Oklahoma Press.

Kohn, H. (1941). *Revolution and dictatorship.* Cambridge, MA: Harvard University Press.

Linz, J. J. (1970). An authoritarian regime: Spain. In E. Allardt & S. Rokkan (Eds.), *Mass politics* (pp. 251–283). New York: Free Press.

Machiavelli, N. (1976). *The discourses.* Middlesex, UK: Pelican Press.

Morlino, L. (1990). Authoritarianism. In A. Bebler & J. Seroka (Eds.), *Contemporary political systems: Classifications and typologies* (pp. 91–119). Boulder, CO: Lynne Rienner.

Perlmutter A. (1981). *The military and politics in modern times.* New Haven, CT: Yale University Press.

Schmitt, C. (1921). *Die diktatur* [The dictatorship]. Munich-Leipzig, Germany: Duncker & Humblot.

DIPLOMACY

In its broadest sense, diplomacy refers to the conduct of human affairs by peaceful means, employing techniques of persuasion and negotiation. In the more specific sphere of international politics, through the utilization of such techniques, it has come to be regarded as one of the key processes characterizing the international system and a defining institution of the system of sovereign states—often referred to as the "Westphalian" system after the 1684 Peace of Westphalia. Its usage, however, embraces some important distinctions. First, at the state level, it has frequently been used (particularly in studies of diplomatic history) as a synonym for *foreign policy*—as in "Russian," "German," and "Japanese" diplomacy (foreign policy). More commonly, however, it is used to refer to one means by which such policies are implemented. Second, viewed as an institution of the international system, a distinction can be made between diplomacy as a set of *processes* and as a set of *structures* through which these processes are conducted. Debates about the continuing utility of diplomacy in contemporary international politics frequently reflect confusion between these meanings. In the course of the following discussion, the origins of diplomacy are outlined, together with differing analytical approaches to its nature and significance as a feature of international politics. The changing nature of diplomatic processes is then discussed, followed by an examination of the evolution of the structures through which diplomacy has been conducted at both the state and international levels.

Theoretical Approaches

While the study of diplomacy has a long and honorable tradition dating back to Machiavellian thought, it is only in recent years that diplomatic practice has started to receive detailed theoretical attention. There is a notable absence of conscious theorizing in much of the scholarship on diplomacy. Instead, the bulk of the scholarship offers detailed historical accounts of diplomatic events (diplomatic history) as well as texts on diplomatic practice. Most scholars of diplomacy implicitly choose from a very narrow range of analytical frameworks drawn almost exclusively from the realist tradition in international relations (IR). As a consequence, the orthodox study of diplomacy has been marked by a remarkably unified theoretical approach—something quite unique in political science. There is a surprising ontological consensus about what diplomacy is and who the diplomats are. This consensus arises from the dominant influence of rationalist thinking. The upshot of this dominance is that the range of the scholarship in a majority of studies of diplomacy tends to be limited to analysis of the international realm of sovereign states in the context of high politics. There is little enthusiasm among mainstream scholars to explore the diplomatic world beyond interstate relations or low politics and that relating to the nonpolitical. This is now changing. There is a growing body of work interested in diplomacy not simply as a foreign policy tool of states but as a means of connecting cultures, polities, economies, and societies. This section aims to highlight briefly the orthodox and unorthodox approaches to the study of diplomacy.

Diplomacy and Realism

The realist core of orthodox approaches to diplomacy is undisputed and is clearly evident in a number of key features found in this prevailing approach. The most telling is the focus on the sovereign state as the primary unit of analysis in diplomacy such that the study of diplomacy is confined to the study of the process and content of interstate relations—that is, how sovereign states seek to engage with each other. Prevailing models of diplomacy focus almost exclusively on singular state-to-state relations. The orthodoxy defines diplomacy as processes of communication, negotiation, and information sharing among sovereign states. Diplomatic processes revolve around the activities of professional diplomats—that is, officials of foreign ministries and overseas missions. More common, especially in North American scholarship, is the narrower definition of diplomacy as a foreign policy tool of states—that is, diplomacy as statecraft. This more limited definition has led to a great deal of foreign policy analysis passing itself off as diplomatic studies, despite the fact that it does not consider the processes of interstate relations as its main focus. Both definitions, however, share the view that diplomacy has an ordering role to play in the otherwise anarchic and unstable international system of states—a view that has theoretical roots in realism. Successful diplomacy, it is argued, creates a system of states. It constructs balances of power, facilitates hegemonic structures, and fashions post-hegemonic regimes. When diplomacy fails or is absent, conflict and war usually follow. Indeed, it is the very fact of conflict between states (a core realist assumption) that warrants the emergence of diplomatic systems.

Beyond Realism

In recent years, however, there have been significant conceptual shifts in the study of diplomacy, and as a result, those studying diplomacy are able to choose from a wider range of analytical approaches. The customary view of what the proper study of diplomacy entails is now contested by scholars who apply analytical strategies drawn from constructivist, postmodern, and critical IR theory to draw attention to the necessity of understanding IR—and diplomacy—beyond the state and the international state system. As a consequence, the study of diplomacy has stepped outside the narrow state-centric security nexus into a world of diplomacy that is more varied but also more difficult to specify. It is perhaps this lack of specificity in what is being analyzed, and why, that explains why unorthodox approaches continue to be marginalized.

Unorthodox approaches are analytically diverse, yet they share a key point of departure from orthodox approaches—a refusal to accept the state as the exclusive unit of diplomatic analysis. Diplomacy is seen as a more open-ended process where diplomatic agency includes not only the state but also a range of nonstate actors such that a sociological concept of diplomacy emerges where diplomacy possesses economic, cultural, social, as well as political forms and functions. A common theme within these approaches is, therefore, the problematic core idea of the foreign ministry and its overseas missions as the sole agent of diplomacy. Unorthodox approaches suggest that the proper terrain of the study of diplomacy includes, and extends beyond, foreign ministries, overseas missions, and the state officials that work in these government institutions and international organizations to potentially include diplomatic networks largely drawn from all sections of domestic and international society covering any number of issues from the environment and e-commerce to avian flu and land mines. An important implication of this is that diplomacy has many modes, including conventional interstate relations, nonconventional intercultural relations or commercial relations, and modes that mix the two. Moreover, the study of diplomacy entails the rejection of the simple reproduction of the status quo of interstate power relations (described as antidiplomacy) at the heart of orthodox studies of diplomacy and, in the case of postmodern approaches, the production of the concept of "otherness," which, it is claimed, is the core of all diplomatic modes. In this sense, the world of diplomacy is characterized not by the commonality of the material and security interests of states but by differences—different interests, diverse cultures, and varied identities.

While orthodox approaches import analytical tools from realist IR to develop concepts such as summit diplomacy, bilateral diplomacy, and multilateral diplomacy, unorthodox approaches, by contrast, import analytical tools from other social science fields such as political economy, business and management studies, philosophy, theology, sociology, and anthropology to explore diplomatic practice. As a result, new concepts of diplomacy such as catalytic diplomacy, network diplomacy, sustainable diplomacy, and multistakeholder diplomacy have been developed to provide analytical means to explain the contemporary diplomatic practice in ways that draw attention to the different interests, cultures, and identities represented by state and nonstate actors and the varied modes of diplomacy that emerge with the dual engagement of this "otherness." These new concepts also draw attention to the changing character of contemporary diplomatic forms, sometimes casting doubt on the notion of the progressive development of diplomatic systems found in traditional approaches to diplomacy.

In sum, unorthodox approaches to diplomacy do not always tie diplomatic practice to the state or to the problem of anarchy. Instead, diplomacy is seen as a means of connecting individuals, groups, societies, economies, and states to build and manage social relations in domestic and systemic environments. By moving beyond traditional realism, unorthodox approaches to the study of diplomacy have promoted greater theoretical reflection and created an intellectual multiplicity in the analysis of diplomatic practices, modes, and processes. It is to these practices, modes, and processes that we now turn.

The Emergence and Development of Diplomacy

As the above discussion of the theories of diplomacy indicates, those who study diplomacy remain divided over whether it is essentially a state-based set of political processes or whether it is a set of network-based political processes. Those who maintain that diplomacy is primarily the pursuit of the foreign policy interests of the state in the international system of states argue that diplomacy is confined to a quite narrow set of bilateral and multilateral processes of communication, representation, and mediation focused on the foreign ministry and its overseas missions. Diplomatic processes continue to exhibit some regularity so that functions, institutions, codes, conventions, and cultures of diplomacy are marked by continuity and marginal change and so that diplomatic rules and norms continue to hold in the future. The obvious casualty in this approach is any in-depth analysis of change in diplomatic structures and processes. By contrast, those who conceptualize diplomacy outside state-centric framework, tend to emphasize continual change in the conduct and context of diplomacy. The principle objective of network-based approaches is to highlight and analyze the challenges posed to diplomacy by contemporary changes in the international system. Scholars turn to issues of globalization and regionalization to emphasize the increasingly complex social, economic, and political context of diplomacy (at domestic, regional, and international levels). For these scholars, change and transformation in diplomatic processes and structures is the central concern of analysis, and in this frame, diplomacy is seen to have both formal and informal structures. Diplomatic processes are network based and draw in a range of public and private actors; there is an absence of agreed rules and norms of diplomatic engagement such that new codes and conventions are emerging or in need of development. In short, diplomacy, in terms of both the varying processes through which it is effected and the machinery through which it is conducted, is a closely linked phenomenon that is the subject of differing interpretations. We will now examine how these have developed in response to changes in both domestic and international environments.

The Origins of Diplomacy

The origins and development of diplomacy are frequently equated with that of the European system of states. In this view, it is associated with the system of states that emerged and consolidated its forms and practices in the wake of the 1648 Peace of Westphalia marking the end of the Thirty Years' War. However, diplomacy and its institutions have a much longer, and more complex, pedigree and have been identified as existing in some of the earliest human societies. Rather than being associated with a specific historical era, diplomacy has been seen as a response to a set of needs and requirements—namely, the mediation of separateness between communities and the desire and need to establish modes of communication between them. Thus, the earliest documents recording what we would now regard as formalized diplomatic practices are to be found in approximately 2500 BCE in what is now the Eastern Mediterranean and the Middle East, and one of the most familiar features of contemporary diplomacy—namely, the practice of using resident ambassadors—predates its usage in the modern European context by some 3,000 years. Elements of diplomacy and diplomatic institutions can be identified in a variety of international systems, including those of Greece, Rome, and China, but none of these possessed what has come to be regarded as the key characteristics of a fully fledged diplomatic system—namely, effective communication, a set of procedures and conventions governing patterns of communication, and a capacity to mediate between diverse cultures. In general terms, the development of diplomacy has been determined by the character of the societies that it has sought to mediate, the international environment, the available modes of communication, and the technologies that determined them.

Thus, in the European context, the medieval era witnessed the growth of diplomatic processes as international relationships became more complex and dense. But this occurred in a period when the sovereign state as we recognize it today had not emerged. Against a background in which universalist ideas represented by the concept of Christendom underpinned by the authority of the Catholic Church and the Holy Roman Empire were a dominant reality, diplomacy was not yet associated with the state, involving diverse political units. Moreover, while rulers engaged in the sending of missions to one another, so too did other entities—commercial, ecclesiastical, and private—in the medieval landscape. Relative distance underscored by the difficulties of communication meant that the dispatch of diplomatic missions was infrequent, and their success marked by a high degree of uncertainty resulting from the hazards associated with medieval travel. Furthermore, the precise functions of the representative were circumscribed. In the early part of the Middle Ages, the most common diplomatic agent was the *nuncius*, whose function was to act as the mouthpiece of the principal on whose behalf he was acting and whose capacity to negotiate was nonexistent or strictly circumscribed. While this matched the requirements of the period in which it developed, the growing complexity of interactions marking the later Middle Ages required the use of officials (procurators) granted the ability to engage in negotiations. In short, while we can see the beginnings of the European diplomatic environment, this was the presovereign-state phase of diplomacy marked by quite fluid and flexible procedures representative of a period of major social, political, and economic change.

It was during the 15th and 16th centuries that a clear outline of the diplomatic system as it was to develop over the next 200 years became visible. By this time, the collapse of the universal concept of Christendom had been accompanied by the gradual emergence of the sovereign state. This not only required a greater capacity to communicate within a changing political and economic environment and thus the development of ways in which this could be effected but also provided enhanced domestic administrative resources necessary to its operations. It was in Northern Italy that the earliest manifestations of this new phase of diplomacy were commonly identified. Here, an early form of what was to become the European system of states could be seen. Significant factors were the geographical proximity of the Italian city-states, their relative similarity in terms of power and thus an inability to exercise hegemonic power, and a shared cultural environment that facilitated communication.

Modern Diplomacy

Against this background, the practices of modern diplomacy were honed. In institutional terms, the key development was the growing utilization of the resident ambassador. As noted above, it was not that this practice was unknown in earlier periods but that diplomacy by mission for specific purposes was far more common, meeting the perceived requirements of the time. Again, it was a combination of political and social change and the consequent requirements imposed on diplomacy that underlay this development. In particular, while its ceremonial and symbolic functions remained significant, a growing need for the gathering of reliable and continuous information replaced the earlier emphasis on the exchange of messages. Gradually, the practices developed in this region of Europe were to spread across the continent and, subsequently, would be adopted as key principles for the conduct of diplomacy as the international system expanded beyond its shores.

These principles assumed several forms. On the one hand, as already seen, more regularized and permanent structures were deemed appropriate and necessary. During the ensuing centuries, the exchange of permanent representatives between national governments would become the norm of diplomatic intercourse as its structures and processes were aligned with the state. Consequently, the rules and norms of diplomacy were refined to support the diplomatic system through the consolidation of the principle of immunity for diplomats and the development of protocol—such as the rules of precedent—established at the 1815 Congress of Vienna and codified in the 1961 Vienna Convention on Diplomatic Relations. The growing importance attached to the practice of diplomacy is reflected in the numerous treatises on the necessary qualities required of the ambassador, notably in the writings of Bernard du Rosier, Philippe de Commynes, and François de Callières. Not only did these writings serve as diplomatic manuals for ambassadors, they reflected the ways in which diplomacy was adapting to the realities of a developing system of sovereign states. At the governmental level, enhanced importance was attached to the capacity to process the growing flow of intelligence generated by diplomatic networks. France, under Cardinal Richelieu, is credited with the creation of the first recognizable foreign ministry in the early 17th century and the gradual separation in the conduct of domestic and foreign policy. Thus, by the 18th century, the patterns of diplomacy at both the international and national levels had assumed the shape that would become a familiar feature of the international order in the ensuing centuries.

The 19th century witnessed the consolidation of these patterns but, at the same time, saw considerable change in response to developments at national and international levels. At the national level, the administrative apparatus for the conduct of diplomacy would become larger, more elaborate, and more professional. This not only reflected changes in the role and structure of the state and the consequent need for more sophisticated bureaucratic systems but also mirrored the growing complexity of foreign policy and the demands that this placed on national governments. The emergence of the modern state and the professionalization of bureaucracy affected foreign ministries, as did the profound social change. Although a gradual and uneven process, the image of diplomacy as the preserve of the aristocracy was weakened as recruitment became less a matter of patronage and more a matter of talent. Across Europe, foreign ministries and their diplomatic services developed systems of recruitment, selection by means of examination, promotion by merit rather than patronage, and embryonic training programs. An early form of the latter was the creation of the Oriental Academy (later Consular Academy) established in the Hapsburg Empire in the mid-18th century. From a focus on language training, diplomatic education was to expand in scope to include aspects such as commercial diplomacy as international economic linkages developed. Despite the fact that diplomacy had begun to embrace the middle classes and to lose some of its aristocratic connotations, the ambience that the latter bestowed on it was slow to disappear, particularly in some European states. In France, for example, by the early 20th century, the diplomatic profession was dominated by the middle class, whereas in Germany, the nobility were a dominant presence. One factor that assisted this process of democratization was the recognition that diplomats could not be expected to finance their activities from their own resources, which had, at least in part, been a feature of past practice, and that unpaid attachés seeking an opening in diplomacy were no longer part of a professional service, hence the development of career structures and the grading of salaries, however meager these might be, alongside. Nevertheless, none of this was to take from diplomacy the air of exclusivity that, to a degree, it continues to possess and that came to be seen as a feature of what would be designated as the "old diplomacy."

Many of the characteristic structures of the foreign ministry would also be established in this period. One of these was the distinction between geographical and functional organizational principles, the latter a recognition of the growing complexity of IR that cut across the division of the world into geographical regions. It would, however, be incorrect to assume that the foreign ministry was a focus of policy making. Much of its work was of a clerical-administrative nature with foreign policy being made at the political level with direct communications between a foreign minister and ambassadors. Moreover, despite later assertions on the part of foreign ministries that they had once been the gatekeeper between states and their international environments, it was not always the case that they enjoyed the privileged position that this implies. For some states, it was other departments—particularly those overseeing commercial relations—that were regarded as possessing greater functionality and prestige.

Developments at the national level went hand in hand with those at the international level. This was represented by the extension of the diplomatic network. By the latter part of the 19th century, all the great powers of Europe had exchanged missions, thus marking bilateral diplomatic relations conducted through permanent residential posts as the hallmark of the international diplomatic system. Furthermore, the practice had spread beyond European shores. Despite suspicions of, and a reluctance to engage in, what was often regarded as a manifestation of old world ills, the United States began to expand its diplomatic service during the 19th century. Elsewhere, countries as diverse as Japan, Persia, and Brazil would develop the makings of a diplomatic machinery at home and a diplomatic service overseas. But of particular significance was the gradual spread of European diplomatic norms—not least those relating to diplomatic privilege and immunities—of great significance as the international system became global in its scope during the 20th century.

Alongside these developments was the emergence of "conference" diplomacy heralding the growth of multilateral diplomacy in the ensuing decades. One manifestation of this was the short-lived "Congress System" following the end of the Napoleonic Wars, comprising sovereigns and chief ministers. Greater effects flowed from the creation of standing conferences of ambassadors in major capitals dealing with specific issues—such as that set up in London on the abolition of the slave trade after the Congress of Vienna. By the latter half of the century, technological developments, particularly in communications, had prompted recognition of the need for international cooperation in areas such as telegraphic (the International Telegraphic Union created in 1865) and postal (Universal Postal Union created in 1874) communications. This not only generated an awareness of the need for diplomatic activity in this area, it also brought with it two effects that would become themes for diplomatic change in later years. One was the need for technical expertise in diplomacy, and the other was the gradual involvement of "domestic" departments in international negotiations, an early example of which was the representation of the Home Office in the British delegation to the 1910 International Aerial Navigation Conference. By the eve of the Great War, then, not only had the structures and processes of diplomacy assumed many of their modern forms, the challenges that they would confront in the coming decades were equally identifiable.

Diplomacy in the 20th Century

Developments in the 20th century posed just as many challenges to diplomacy as those in the previous century and indeed some developments were to exact a high price on the reputation and prestige of the diplomatic system and professional diplomats. The crisis of the Great War (1914–1918) was one such development, and it holds particular significance since it led to widespread condemnation of the old European-based diplomacy, which had not only failed to prevent war but also had, as many concluded, contributed to its outbreak. One key failing of diplomacy in this period was the abandonment of established diplomatic channels of communication by resident ambassadors in favor of more secretive diplomatic practices. Diplomacy quickly deteriorated into a closed system, where behind-the-scenes bilateralism and propaganda fed an appetite for the brutal pursuit of national interest in an atmosphere of mistrust and rivalry that the diplomats themselves had done much to create. The old diplomacy that emerged at the end of the 19th century had created the very problems that drove the European powers to all-out war within the first 2 decades of the following century, and condemnation of the old diplomacy quickly led to demands for a new, and more open, diplomatic system where diplomats could be held accountable to their executives.

European states' response to these demands was almost universal. Across Europe, the semiautonomous resident ambassador was replaced by a centrally controlled system of overseas permanent missions. While bilateralism remained a core diplomatic process around these new permanent overseas missions, during the Great War period and in the following decade, multilateral diplomacy took off. A significant amount of intergovernmental diplomacy was now taking place outside the more established bilateral diplomatic structures and foreign ministries in the form of intergovernmental conferences of state leaders and other government ministers such as finance and trade. Multilateral diplomacy involving officials from departments across government became a key vehicle for allied cooperation during the war on issues such as food and munitions transportation, as well as intelligence sharing and military coordination.

Important and influential though it certainly was, the Great War did not exhaust the challenges that the so-called new diplomacy would face in the 20th century. Within just a few decades, the European-based diplomatic system would be both overhauled and expanded to other continents as the diplomatic system adapted to two open-ended developments: (1) the growing *interdependence* of states in the international system, which increased the demand for effective coordination of international cooperation in an ever-growing number of policy areas but especially in trade and finance, and (2) the onset of *decolonization and independence* that more than quadrupled the number of sovereign states in the international system by the end of the 20th century.

The Great Depression of the 1930s demonstrated very starkly the economic interdependence of states, and after the end of World War II, financial and trade integration intensified, creating demands on, as well as opportunities for, diplomats to coordinate international economic policy in bilateral, regional, and multilateral relationships. Indeed, the development of the Bretton Woods system created a number of powerful international economic organizations such as the General Agreement on Tariffs and Trade where diplomats would negotiate international trade policies and rules governing trade. Economic interdependence within regions led to a huge expansion in the number of regional organizations in Europe, Asia, the Americas, and eventually Africa and the Caribbean. The vast increase in the number of international organizations from around 40 at the beginning of the 20th century to almost 400 at its end is one of the most significant developments in IR during this period. These organizations would also present both opportunities for, and demands on, diplomacy to work in new institutional and policy environments and to develop new diplomatic methods in, for example, multilateral bargaining. With the development of nuclear arms and the Cold War from the 1950s and international terrorism from the 1970s onward, strategic interdependence between states became worldwide. Diplomats would populate the increasing number of multilateral and regional strategic organizations from the United Nations Security Council and the International Atomic Energy Agency to the North Atlantic Treaty Organization and the Warsaw Pact. Diplomats would quickly develop new diplomatic methods in, for example, coercive diplomacy, deterrence, and intelligence gathering, as well as management of new processes such as superpower summitry and international peacekeeping. The development of rapid mass communications in the 20th century linked domestic developments in one country directly with others, again creating opportunities for, and demands on, diplomats to develop new practices in order to influence political and policy developments in each others' countries. These included public diplomacy. Interdependence essentially brought an international dimension to almost all aspects of policy and strategy such that the realm and content of diplomacy during the 20th century covered the entirety of world governance.

While interdependence between societies intensified and deepened as a result of globalization and regionalization processes throughout the 20th century, it was especially so in the latter decades as new technologies expedited and reduced communication and transportation costs. The pattern of IR would change dramatically during this period as transnational relations between nonstate actors developed and new global and regional actors from the private sector and civil society emerged and sought to influence policy and processes at all levels creating demands for open and accountable global and regional governance processes. These developments posed fundamental questions about the role and influence of the state—and hence diplomacy—relative to other actors in the system such as transnational business and global civil society. If the state was indeed in decline, then state-based diplomatic systems would, it seems, have decreasing utility.

These new patterns—whether viewed as the cause or response—went hand in hand with the development of linkage of increasingly complex and technical policy issues. Issues of interdependence, globalization, and regionalization raised practical matters crucial to the continued effective practice of diplomacy. How could the diplomat—a generalist by nature and by training—have sufficient grasp of such highly technical policy issues and how could the foreign ministry continue to manage policy issues that cut across several domestic department concerns?

At the same time that these pressures were raising questions about the effectiveness of the state and state-based diplomacy, the processes of decolonization and independence in the mid- and late 20th century highlighted the continued relevance of the state and the continuing appeal of European diplomatic institutions. Decolonization and independence of the colonial states and the former Soviet states increased the number of sovereign states in the international system and, almost without fail, each of the new states created diplomatic institutions in the image of the European model of a foreign ministry and system of overseas missions and permanent delegations. The expansion in the number of sovereign states raised questions about the diplomatic capacity of the new states, the impact a flush of new states would have on multilateral and regional diplomacy, and the development of bilateral diplomacy between the new states and the old states.

With the exception of India, which had for some time before independence in 1947 acquired a quasi-diplomatic system to represent itself in, for example, the League of Nations, most new states, and in particular African states, had very limited resources to spend on developing a European-style diplomatic system of an extensive network of overseas missions. Most relied on a handful of diplomats in key international organizations and key capitals, and it became common practice for diplomats from developing countries to provide diplomatic representation in multiple arenas. An African diplomat is often, for example, the permanent delegate to the UN as well as ambassador in Washington and Ottawa. Similarly, the ambassador for Tonga usually fills a number of posts: permanent delegate at the World Trade Organization (WTO) in Geneva and European Union (EU) in Brussels, as well as ambassador in London.

The extent of the newly developed states' involvement in multilateral diplomacy was, inevitably, limited, at least until developing countries began to form strategic alliances such as the Non-Aligned Movement and the Group of 77 in the UN. Newly independent states also created their own intergovernmental organizations, such as the Organization for African Unity created in 1963, as a way of managing regional integration and security as well as building collective diplomatic strategies.

The arrival of many new states into the system raised strategic issues for the old states, since they did not always have, or wish to spend, the resources required to establish diplomatic representation in so many new countries. Internal debates ensued about how to keep costs down and retain effective coverage of key strategic countries and regions. Many European governments, for example, faced demands to reduce bilateral missions in European capitals in the wake of the development of very large permanent delegations to the EU.

Contemporary Trends in Diplomacy

Three broad trends in contemporary diplomacy are now evident: (1) *fragmentation*, as the conduct of diplomacy at the governmental level now involves government departments traditionally associated with purely "domestic" issues; (2) *concentration*, as the fusion of domestic and international politics has been accompanied by the expanding involvement of heads of government in international policy; and (3) *diffusion*, as professional diplomats have found themselves required to engage with a growing range of nongovernmental stakeholders in complex policy networks.

The first of these trends, fragmentation, came to be associated from the 1970s onward with the development of an expanded "foreign policy community"—that is, an expanding range of governmental agencies and a multiplicity of channels in the conduct of external relations supplementing and often challenging the role claimed by foreign ministries. One consequence of this has been a growing emphasis on the need for policy coordination at the national level, underpinned by recognition that an uncoordinated stance in international negotiations reflecting various bureaucratic interests has potential costs in terms of attainment of policy goals. This partly explains the trend toward concentration. Awareness of the potential costs of lack of bureaucratic and political coordination and politicization of international policy, combined with a growing international role for heads of government, have resulted in a tendency to centralize the conduct of diplomacy in, for example, prime ministerial and presidential offices. Additionally, the imperatives of coordination have resulted in the merging of departments in the quest for greater efficiency in the management of external relations, notably in the area of external trade relations. Hence, both Canada and Australia merged their foreign ministries and international trade departments during the 1980s.

The third trend, diffusion, reflects the fact that diplomatic processes have increasingly required the development of policy networks as complex policy issues demand that state-focused diplomacy is supplemented by linkages with civil society organizations and, in specific contexts, the business community. Part of this development mirrors the changing nature of international negotiation as it assumes the character of a management process marked by its technical qualities, complexity, uncertainty, and bureaucratization. In this context, diplomacy has become much more than the trading of concessions in pursuit of a negotiated settlement. Many contemporary negotiations, such as those in the area of the environment, involve processes of mutual learning and the creation and systematizing of new knowledge and mutual education among a group of interests, each of which has contributions to make to the management of policy issues. The impact of this can be seen not only in the growing engagement between actors at the national level but also in multilateral diplomatic environments. Stakeholder engagement has become a watchword in the majority of international organizations, from the UN to the World Bank and the WTO. Underpinning these developments at the national and international levels is the enhanced emphasis on public diplomacy from the late 1990s onward. But from being primarily concerned with image management, public diplomacy strategies are increasingly founded on an awareness that the routes of influence within the international system are changing and that shaping international policy outcomes demand strategies for influencing a much wider range of constituencies than those at the governmental level. Consequently, by the early 21st century, public diplomacy had become one of the primary concerns of foreign ministries. One obvious manifestation of the awareness of the importance of engaging with an expanded audience at home and abroad has been the attention paid by foreign ministries to the use of the Internet and increasingly sophisticated and accessible Web sites. Recognition of the growing demand for interactive processes means that these ministries are now using social networking sites such as Facebook together with YouTube and Twitter to engage in a dialogue with audiences at home and overseas. Hesitant forays into the virtual world, Second Life, have begun. Sweden has established an embassy there and the U.S. State Department has sought to develop its public diplomacy strategies by using this resource to engage with bloggers in the Middle East.

Change has not been limited to the foreign ministry, however. The structure and operations of diplomatic services have undergone significant change, partly as a result of resource constraints but also due to the evolving international order itself. Here, the need to operate more economically has combined with an awareness that patterns of international representation have failed to keep pace with geopolitical and geoeconomic change. In the United States, the announcement by former Secretary of State Condoleezza Rice of the Transformational Diplomacy program was stimulated by these concerns. In addition to this, the impact of the twin processes of fragmentation and consolidation has meant that the overseas network of diplomatic posts is now regarded as representative of the entire governmental apparatus and not only of the foreign ministry. Indeed, in some diplomatic posts, the number of traditional diplomats is dwarfed by the presence of officials from a range of "domestic" government departments.

A further significant development is the revolution in communications and the role of the media in international policy. The adoption of e-mail and secure facsimile links between foreign ministries and overseas missions (contrary to the traditional arguments concerning the impact of enhanced communications on diplomacy) has allowed missions to play a more direct role in the policy processes. Simultaneously, the development of the electronic mass media creates pressure on governments to respond to events almost instantaneously and, at the same time, provides opportunities to project their policies to domestic and foreign audiences. Again, this has implications for the respective roles of foreign ministries and diplomats in the field. The enhanced speed of events can often assume as great an importance as the events themselves. Nowhere is this more evident than in natural and man-made disasters. The growing incidence of terrorist attacks and events such as the tsunami in 2004 has placed renewed emphasis on the consular dimension of diplomacy. The rise of mass tourism and a media ready to judge diplomats by the immediacy of their response to such crises has established a new benchmark by which the diplomatic profession is judged.

Cumulatively, these developments have had a significant impact on the role of the professional diplomat. Increasingly, this is portrayed in terms of a "coordinator-manager" and "facilitator" in complex processes spanning the boundaries between the international and national domains. They also pose questions concerning the traditional norms of behavior associated with diplomacy. To take one significant example, the traditional emphasis on confidentiality and secrecy is challenged by the norms demanded by the need to work with a range of "nondiplomatic" stakeholders in specific policy milieus. The latter, working to their own codes and norms of behavior, do not always respect the traditions associated with the conduct of diplomacy, and establishing mutual understanding and cooperation is one of the major challenges of contemporary diplomacy.

Adaptation and Change in Diplomacy

Several clear themes emerge from this discussion of the emergence and development of diplomacy. The first is that diplomacy has a history and a logic that transcends the system of states with which it is often equated. Just as diplomacy preceded the emergence of the sovereign state, so also it has adapted to the latter's transformation in response to forces associated with globalization and regionalization. The second theme relates to the first: Diplomacy has a capacity to adapt to change. It is something of a truism—and a recurrent topic of diplomatic memoirs—that diplomacy is not what it was. Thus, change runs as a leitmotif through its evolution as the processes of diplomacy and the structures through which it has been conducted have responded to transforming environments. This, however, has been a gradual process, so much so that the frequently used distinction between "old" and "new" diplomacies is misleading, ignoring as it does the inherent adaptive capacity of the processes on which diplomacy has relied over time. Similarly, debates about the "decline" of diplomacy are usually founded on the association of these processes with particular structures or forms—such as the emergence of resident bilateral diplomacy. To understand the nature and significance of diplomacy in its historical and contemporary manifestations we must recognize that there is no single mechanism through which its objectives can be served. Diplomacy's capacity for change ensures that in the early 21st century, where complex agendas require evermore inventive modes of global governance, the processes associated with diplomacy remain a major component of international life.

Brian Hocking
Loughborough University
Leicestershire, United Kingdom

Donna Lee
University of Birmingham
Birmingham, United Kingdom

See also Anarchy; Balance of Power; Bilateralism; Economic Statecraft; Foreign Policy Analysis; Globalization; Governance, Global; International Organizations; International Society; International System; Mediation in International Relations; Multilateralism; Networks; Nonstate Actors; Policy Network; Realism in International Relations; Regionalization; State; United Nations; War and Peace; Westphalian Ideal State; World Bank; World Trade Organization (WTO)

Further Readings

Bátora, J. (2008). *Foreign ministries and the information revolution: Going virtual?* Leiden, Netherlands: Brill.

Berridge, G. R. (2005). *Diplomacy: Theory and practice* (3rd ed.). Basingstoke, UK: Palgrave.

Cohen, R., & Westbrook, R. (Eds.). (2000). *Amarna diplomacy: The beginnings of international relations*. Baltimore: Johns Hopkins University Press.

Cooper, A., Hocking, B., & Maley, W. (Eds.). (2008). *Global governance and diplomacy: Worlds apart?* Houndmills, UK: Palgrave.

Der Derian, J. (1987). *On diplomacy: A genealogy of Western estrangement*. Oxford, UK: Blackwell.

Hamilton, K., & Langhorne, R. (2000). *The practice of diplomacy*. London: Routledge.

Hocking, B. (Ed.). (1999). *Foreign ministries: Change and adaptation*. London: Macmillan.

Jönsson, C., & Hall, M. (2005). *Essence of diplomacy*. Houndmills, UK: Palgrave.

Lee, D., & Hudson, D. (2004). The old and new significance of political economy in diplomacy. *Review of International Studies, 30,* 343–360.

Mattingly, G. (1973). *Renaissance diplomacy*. London: Penguin.

Melissen, J. (Ed.). (1999). *Innovation in diplomatic practice*. London: Macmillan.

Neumann, I. (2002). *The English school on diplomacy* (Discussion Papers in Diplomacy No. 79). The Hague, Netherlands: Netherlands Institute of International Relations.

Nickles, D. P. (2002). *Under the wire: How the telegraph changed diplomacy*. Cambridge, MA: Harvard University Press.

Nicolson, H. (1963). *Diplomacy*. Oxford, UK: Oxford University Press.

Rana, K. (2000). *Inside diplomacy*. New Delhi, India: Manas.

Sharp, P. (1999). For diplomacy: Representation and the study of international relations. *International Studies Review, 1,* 33–57.

Watson, A. (1982). *Diplomacy: The dialogue between states*. London: Eyre Methuen.

DISARMAMENT

The term *disarmament* is generally used to describe a process of reducing and eliminating certain weapons systems, but it may also depict an end state when a specific type of weapon has been abolished. The quest for a definition is complicated because the concept of disarmament may be employed for different purposes in diplomacy, in international relations, and in national and international security debates. Disarmament can apply to any armaments system, from nuclear weapons to land mines, and depending on political perspective may be cast as negative (inducing vulnerability) or positive (promoting peace).

Though disarmament may be linked with arms control, the two approaches are different. Arms control covers partial measures and mechanisms for managing and restricting the development and possession of certain kinds of armaments, whereas disarmament is the process or accomplishment of the elimination and abolition of such weapons systems. Arms control may be an end in itself, leaving some weapons in the hands of some countries, or it can provide a stepping stone toward disarmament, but the two concepts are not interchangeable and should not be confused. Concepts of disarmament are related not only to reducing and eliminating weapons but also to challenging traditional, military concepts of security, defense, and deterrence and constructing alternative approaches for peace and security. This entry discusses various approaches to disarmament.

General and Complete Disarmament

When used in relation to nuclear weapons, disarmament and nonproliferation are frequently described as two sides of the same coin, implying that one is not sustainable without the other. The Treaty on the Non-Proliferation of Nuclear Weapons (NPT), which entered into force in 1970 with the purpose of preventing the spread of nuclear weapons and technologies, contained in its Article VI an obligation to pursue nuclear disarmament. Article VI also referred to "general and complete disarmament," a phrase drawn from United Nations (UN) resolutions going back to 1946. No timetable was set for achieving these objectives, and there are many competing interpretations of what would actually need to be done to implement Article VI.

In the aftermath of World War II, "general and complete disarmament" reflected the aspiration of resolving all conflict through diplomacy, without recourse to force and arms, but its meaning has never been clarified. Was it intended to mean a world without all weapons or a world without nationally controlled war-fighting weapons? It appears that the founders of the UN wanted this aspiration to guide international relations toward more peaceful directions. Certainly, the world would have to change fundamentally if total disarmament, peace, and security were ever to be possible. The inclusion of "general and complete disarmament" in Article VI of the NPT has more often been used by some of the nuclear weapon states (defined in the NPT to include China, France, Russia, the United Kingdom [UK], and the United States) to justify not eliminating their nuclear weapons until there is general and complete disarmament. So that the laudable but remote goal of total disarmament would not be used as an excuse to hold on to nuclear weapons, the Sixth Review Conference of the NPT, held in 2000, made clear in its consensus final document that nuclear disarmament is the most urgent priority under the treaty and that achieving nuclear disarmament is not contingent on general and complete disarmament, which is the "ultimate goal."

Disarmament Initiatives, Conferences, and Agreements

While the aspiration of general and complete disarmament is unlikely ever to be fully achieved, much can be done to reduce and eliminate many weapons systems. From government-led action to remove all guns from Japan in the 15th and 16th centuries to negotiations on banning cluster munitions in 2008, there have been a range of disarmament initiatives, conferences, and agreements. The Hague Peace Conferences that took place from 1899 effectively outlawed dum-dum bullets. After World War I, the Geneva Conventions were developed to regulate the conduct and customs of warfare, including aspects of arms developments and trade. The 1925 Geneva Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous, or Other Gases, and of Bacteriological Methods of Warfare (known as the Geneva Gas Protocol) was an early disarmament effort that focused on banning the *use* of inhumane weapons rather than mandating physical disarmament of the actual weapons. It took a further 47 years to negotiate the 1972 Biological and Toxin Weapons Convention, which prohibited the development, manufacture, and stockpiling of bioweapons.

During the Cold War, arms control was considered to be an important part of the strategic relationship between the United States and the Soviet Union. This was particularly true of nuclear weapons, where a number of bilateral agreements sought to cap a spiraling arms race in which the dominant powers developed more than 70,000 nuclear weapons. One of these, the 1987 Intermediate-Range Nuclear Forces Treaty, removed and eliminated all nuclear-armed Cruise and Pershing Missiles from Europe. This treaty was brought about by changing perspectives and leadership in the Soviet Union and the United States in response to a massive upsurge of civil society protests and campaigning for nuclear disarmament and human rights in the 1980s. Unusually, it met the criteria for both disarmament and arms control.

In political discourse, disarmament is often qualified as "unilateral," denoting one country's decisions and actions to renounce a particular type of weapon, and "multilateral," denoting decisions and processes undertaken by many states, generally through negotiations. Though unilateral and multilateral approaches may be different, they are not mutually exclusive and can be mutually reinforcing. Nation-states are members of the various bodies and institutions that constitute the world's disarmament machinery. The Conference on Disarmament (CD) was formally constituted in Geneva following the First UN Special Session on Disarmament in 1978, which had identified the need for a "single multilateral disarmament negotiating forum of limited size taking decisions on the basis of consensus." The CD accomplished little until the Cold War ended. In 1993, it concluded negotiations on the Chemical Weapons Convention, which prohibited the acquisition, manufacture, and development of chemical weapons and created the multilateral Organization for the Prohibition of Chemical Weapons to oversee and verify the complete elimination of such armaments. As its membership went from 38 to 66 UN member states, the CD negotiated the Comprehensive Test Ban Treaty (CTBT), which was concluded and opened for signature in 1996. By banning nuclear tests, the CTBT contributes to nuclear disarmament as well as nonproliferation.

The UN General Assembly's First Committee covers international security and disarmament issues and meets annually to consider between 60 and 70 resolutions on these issues, ranging from nuclear weapon–free zones to prevention of an arms race in outer space. There is also a UN Disarmament Commission, which is open to all UN member states, but this meets only for 2 weeks every year and does not have decision-making powers. The UN Office for Disarmament Affairs is headed by the Secretary-General's High Representative for Disarmament Affairs and provides support for multilateral treaties and UN-related institutions dealing with weapons of mass destruction and conventional arms, including small arms and light weapons.

At the end of the Cold War, disarmament was promoted for humanitarian as well as strategic or conflict reduction purposes. In the field of conventional weapons, civil society—notably disarmament campaigners, war veterans, and humanitarian aid workers—built momentum around the demand to ban land mines on grounds that the weapons continued to maim and kill noncombatants long after the military purpose had ceased. They promoted conditions for a group of concerned middle-power states to establish a forum and begin negotiations to ban antipersonnel land mines. The 1997 Mine Ban Treaty was negotiated outside the traditional multilateral machinery because efforts to get negotiations on this issue in the CD or among parties to the 1981 Inhumane Weapons Convention were blocked by states with a vested interest in retaining the weapons. A decade later, another coalition of governments and civil society created their own series of negotiating conferences to achieve a ban on the use and manufacture of cluster munitions. Though some of the major governments chose to boycott the negotiations, they were not able to block them. The resulting disarmament treaties have proved significant in stigmatizing certain weapons as inhumane and creating a normative framework within which even nonsignatories find that their deployment and use of such weapons is constrained. Once the treaties establish the disarmament norm, they change states' behavior. Most of the states that boycotted the negotiations have later signed and ratified.

Public Involvement

In recent history, it has taken active public engagement and civil society pressure to persuade governments to renounce certain kinds of weapons. Though lip service is paid to the desirability of disarmament, it does not occur in a vacuum. As one type of weapon is banned and disarmed, there are vested interests keen on developing new weapons. Arms trading is big business, netting huge profits for arms manufacturers, defense contractors, governments, and black marketers. Governments may support nonproliferation to disarm others while retaining high-value weapons for themselves.

Where traditional concepts of defense and arms control are associated with national security paradigms, disarmament has come to be associated with the evolving human security paradigm and linked with development, human rights, and a redistribution of resources away from military methods and expenditure.

Considered as an end state, disarmament may never be complete, though specific weapons systems have been comprehensively prohibited and eliminated. Viewed as a process, disarmament is taking place on many levels all around the world: from the women who help implement the Programme of Action on Small Arms and Light Weapons by coordinating local initiatives to collecting and destroying guns to the diplomats negotiating the next nuclear treaties and from disarmament activists halting nuclear weapon convoys to the international inspectors verifying that states are complying fully with their treaty obligations.

Rebecca E. Johnson
Acronym Institute for Disarmament Diplomacy
London, United Kingdom

See also Arms Race; Conflict Resolution; Multilateralism; Nongovernmental Organizations (NGOs); Security and Defense Policy; Security Cooperation; United Nations

Further Readings

Borrie, J., & Randon, V. M. (Eds.). (2006). *Thinking outside the box in multilateral disarmament and arms control negotiations*. New York: United Nations.

Cameron, M. A., Lawson, R. J., & Tomlin, B. W. (Eds.). (1998). *To walk without fear: The global movement to ban landmines*. Toronto, ON, Canada: Oxford University Press.

Goldblat, J. (2002). *Arms control: The new guide to negotiations and agreements*. London: Sage.

Johnson, R. (2009). *Unfinished business: The negotiation of the CTBT and the end of nuclear testing*. New York: United Nations.

Myrdal, A. (1977). *The game of disarmament: How the United States and Russia run the arms race*. Manchester, UK: Manchester University Press.

Discipline of Political Science, The

See *Introduction to Political Science* (Volume 1)

Discontent

See Dissatisfaction, Political

Discourse Analysis

The aim of discourse analysis is to reveal the ontological and epistemological premises that are embedded in language and that allow a statement to be understood as rational or interpreted as meaningful. Discourse analysis investigates whether—in statements or texts—it is possible to establish any regularity in not only the objects that are discussed, the subjects designated as actors, and the causal relations claimed to exist between objects (explanans) and subjects (explanandum) but also the expected outcome of subjects trying to influence objects, the goal of their action, and finally the time dimension by which these relations are framed. Discourses thus comprise the underlying conditions for a statement to be interpreted as meaningful and rational. At the same time, discourse analysis is the study of rationality and how it is expressed in a particular historical context. Discourse analysis is part of the constructivist (or social-constructivist) approach within the humanities and social sciences. It assumes that basic assumptions with regard to being, self, and the world are constructed by individuals living in a historical and cultural context that is produced and reproduced by their speech and acts.

There is no mainstream definition of discourse within the social sciences. Neither is there any generally accepted understanding of what discourse analysis is or which method(s) its practitioners should use. Consequently, it is difficult to give a precise description of what characterizes discourse analysis. This entry reviews several forms of discourse analysis and their application to politics. Three approaches are distinguished, all of which are called discourse analysis, but they alternate in their approaches to what a discourse is and what the aim of analyzing discourses is. The first approach is the discourse-analytical, the second is the discourse-theoretical, and the third is the critical discourse analysis.

Varieties of Discourse Analysis

Analyses of discourse have been carried out within a variety of social science disciplines, including linguistics, anthropology, sociology, international relations, communication studies, and political science. Although the concept of political discourse has been used for centuries to describe political debate or for deliberation in political theory and philosophy, it is only during the past 40 years or so that there has been a theoretical and methodological interest in how to study the relationship between language and political action. This started in the 1960s in Europe as part of a philosophical renewal of the humanities (including the social sciences), later to be known as structuralism and poststructuralism or in more general terms as the linguistic turn. In the 1970s, it spread to the United States with studies of how political concepts and political news play a role in the construction of social problems. Today, there are several approaches on how to understand the role of language in politics. Among these are conceptual historiography (*Begriffsgeschichte*), the history of political ideas, and the theory of narration. They all differ from discourse analysis by the fact that their object of study is concepts, narration, and ideas and not discourses. The most important difference among discourse analytical approaches is between those that seek to understand discourse as a contingent form of knowledge and use discourse analysis to see how knowledge and the production of knowledge have changed over time and those that take for granted that "the world" is a product of how we categorize it through our statements and therefore look on discourse as a universal type of social action and use discourse analysis to establish a general theory of discourse. Although discourse analytic approaches emphasize the connection between discourse and power, they differ in how they attach the concept of discourse to other concepts such as knowledge, ideology, ideas, and truth.

Discourse Analysis

This understanding of discourse and *discourse analysis* is closely connected to Michel Foucault and his publications from 1963 to 1971. In this period, Foucault studied the history of language and how words (or language) were placed in relation to things (or what is observable) at various periods in history. Foucault was concerned with the fact that from the mid-1800s, the human sciences had begun to analyze language and to argue that all human actions and social formations are somehow related to language or can even be understood as constructed in (or by the use of) language. According to Foucault, then, discourse analysis is not some independent theory or method but a way in which the human sciences perceive the world. Foucault shows how language was turned into an empirical object for scientific studies and views discourse analysis as a historically specific manner in which the human sciences relate to reality. Hence, Foucault considers that scientific interest in language is a historical event: the end of the modernity period, where man stood at the center of scientific interest, and the start of a new period, where language became the central object of study. It is in this context that Foucault introduces the concepts of *archives* and *archaeology* in an attempt to portray discourses as historically determined forms of knowledge, which, together with other discourses, enters into a form of institutionalized rationality (an archive). The archive, in consequence, is a historically determined knowledge horizon, a framework for how ideas are produced and sustained and for how knowledge is accepted (as being trustworthy) or not. Ideas are created in discursive events, which subsequently—by historical analysis—can be understood to have added new positions to the archive or to have transformed already existing positions in the archive. At the same time, the archaeology is the knowledge we possess about the history of the various forms of knowledge and of the limitations and possibilities that exist for creating knowledge and generating ideas. These limitations and possibilities are not exclusively linguistic. They are also extradiscursive and institutional. Discourses are supported by institutions, and together with various technologies (e.g., disciplining or sanctioning), they constitute a historically determined rationality.

Discourse Theory

In contrast to Foucault, *discourse theory* aims at developing a universal theory of discourses. Discourse theory sees all social phenomena as discursive constructions and assumes that all social phenomena can be studied by discourse analysis. It is in this sense that discourse theory turns social phenomena into language and language into an object for discourse analysis. Ernesto Laclau and Chantal Mouffe are the best examples. They have advanced discourse theory by deconstructing other theories. First, with inspiration from structuralist linguistics, they emphasize that the smallest unit in a discourse (langue) is the sign and that discourses include a system of signs characterized by every sign being different from other signs. Second, with inspiration from poststructuralism, they emphasize that signs are infused with meaning through articulation (*signifié*), while the content of signs (*significant*) is always contingent and never fixed. Finally, with inspiration from Neo-Marxism, they stress that the articulation is embedded in a political process. In their definition, discourse is a system of signs that allocates meaning through articulation. The articulation, on the one hand, is understood as a conflict between persons whose object is to achieve political status by imposing a particular taken-for-granted understanding of the world. On the other hand, discourse analysis is used to map or trace this process as a political process. The task of discourse analysis is to find the nodal points that give other signs their meaning and to observe the process through which the allocation of meaning is taking place. In political theory, for example, "democracy" is a nodal point around which conflicts are constantly taking place. In contrast to Foucault's discourse analysis, the concept of ideology (or objectivity) plays an important role in discourse theory. All discourses are ideological because they appear as objectivity–that which is taken for granted—and thus conceal alternative realities. Also, in contrast to Foucault, the concept of knowledge does not enter into the vocabulary of discourse theory. Where Foucault can study how knowledge *has become* an archive with his archaeological (diachronic) approach, the discourse theoretical (synchronic) approach is analyzing how meaning *is created* by politics. And finally, where Foucault sees institutions as supporting knowledge and therefore capable of having an independent (nondiscursive) status, in discourse theory institutions are understood as discursive constructions without any extradiscursive status.

Critical Discourse Analysis

Critical discourse analysis emphasizes the necessity of establishing methods for empirical investigation of relations between discursive and nondiscursive practices. In this sense, it distinguishes itself mainly from discourse theory. The work of Norman Fairclough is central here. According to Fairclough, discourse is not only a communicative act but also a social practice. Discourses not only constitute social phenomena but are also constituted by social phenomena in the form of social (or political) practice. Any use of language (a communicative action) therefore consists of a discursive practice where discourses are produced or consumed and a social practice or an institutional context of which a communicative action is a part. The communicative action can draw on (consume) or create (produce) discourses, but it will always be part of an order of discourse, where several discourses are articulated simultaneously. The communicative act is linked to social practice through the use of genres or conventional text types; the news media can, for example, draw on interviews as a genre, while the family can use the dinner table conversation in the same manner. Fairclough combines linguistic textual analysis with macro- and microsociological analysis of texts and conversations, using a comprehensive research design that recognizes five components (problem formulation, choice of empirical methods, transcription, analysis, and results), each extended by specific methods and checklists.

In contrast to the theory of discourse, the critical discourse analysis distinguishes between discourse and institutions as two different types of social phenomena. It studies how discourse and institution interact in the constitution of a social world and how discursive practices are institutionalized or are moved from being linguistic utterances to set conditions for stable social relations. While critical discourse analysis attempts to uncover the ideologies that contribute to the production and reproduction of power, it also has a political aim: It looks not only for how discourses limit our understanding of the world (i.e., function as an ideology) but also for how they contain several competing discourses and therefore the possibility of dominant ideologies being contested. The neo-Marxist concept of hegemony is used in this context. Ideology is understood to be embedded in discursive practice and discourses to be more or less ideological, where the ideological discourses are those that contribute to maintaining (or establishing) a power relation.

Discourse Analytical Approaches to Politics

Discourse analysis has developed rather rapidly from emphasizing the ontological (constructivist or antirealistic) approach to building up methods and tools for studying language and texts. Most recently, the question of method has come to occupy a central position, as has the requirement of being able to evaluate the validity of findings produced by discourse analysis. Still, discourse analysis remains an alternative—critical—approach to mainstream political science and other social sciences. It assumes that political science, like other social science disciplines, is a form of knowledge that constitutes a historically constructed understanding of what can be studied as politics (an ontology) and a historically constructed interpretation of what are assumed to be true statements about politics (an epistemology). It reflects on political science as a discipline and on the role of the discipline in constructing a certain form of political order (or polity). It finally applies concepts such as polity, politics, and policy in ways different from mainstream political science. In contrast to the realism that characterizes mainstream political science, discourse analysis is built on an antirealistic, or constructivist, approach, perceiving any form of political order to be embedded in language (and institutions) and articulated by the use of speech acts. Thus, for discourse analysis, the conflicts through which what *politics is* (and what is nonpolitical) is defined, therefore, are an important dimension to every type of polity and thus also an important object for political science to study.

The most important contribution of discourse analysis to political science lies in the theories of power. From Thomas Hobbes to Robert A. Dahl, power has been seen as the ability to affect, to limit, or to control the behavior of people or as the ability A has to compel B to do something that B would not otherwise have done. In contrast to this view, discourse analysis understands power as entrenched in the taken-for-granted. Foucault speaks about the productivity of power and perceives power as one of the technologies by which individual and collective identities are constructed and in which the understanding of self and of communality comes to be taken-for-granted or to be understood as meaningful and rational. Where mainstream power theory (e.g., Dahl) takes for granted that A and B exist, each with its own set of preferences, Foucault sees the constitution of subjectivity and of the relation between A and B as the productivity of power. His analysis—together with the two other types of discourse analysis—looks for the "hidden power" embedded in both individual and collective identities and for the underlying conditions for interests, expectations, and interpretations to be understood as rational. Apart from regarding power as ontological, the three approaches therefore agree that it is also epistemological. Power is not only the ability to affect the behavior of others. It is also the productive force by which A and B are constructed, each with its set of interests and each with its set of expectations and interpretations.

By combining the critical with the analytical approach, discourse analysis is characterized as standing at the crossroads of several approaches in the study of politics. It opens avenues that bring political science into close relationship not only with history and institutional theory but also with linguistic and narrative theory. Underlying the three discourse-analytical approaches are two approaches to time and space. The first is the diachronic, where the question is raised about how epochs come into being and what, over their long history (the *longue durée*), makes the various types of political orders different from each other. The diachronic understanding generates a history of polities, their periods, and geography, which also allows discourse analysis to sit at the crossroads of history as a scientific discipline (archaeology) and political science as the history of political orders.

Foucault is an example in point. He describes the history of bio-politics, and by doing so also establishes the historiography of a specific type of political technique called governance. The second is the synchronic understanding of time, where the question is raised about how political orders are constructed and changed by the means of politics. By studying how politics constructs political orders, the short-term and long-term perspectives are connected, and the history of political orders is combined in the study of how polities are constructed through changes in understandings of what politics is. The combination of the two makes it possible to study the history of political orders as a precondition for the study of political change. Fairclough is an example in point. He describes how an ideology becomes dominant and how social conflicts are taking place within the context of dominant ideologies. He also studies how discursive practices are institutionalized and how they proceed within a context of institutions. The synchronic process tracking thus generates insight into the political processes as conflicts over dominance, which positions the discourse analysis at the crossroads of institutional theory (and theories of institutional change) and political science as the theory of politics (or political change).

Within the past few years, institutional theory has become aware of the role of ideas in explaining institutional change. Attempts are being made to understand how ideas are causally powerful in explaining the form and content of institutional change. The ideational account of institutional change is thus moving institutional theory in the direction of discourse analysis. This applies to studies of economic ideas and their role in defining economic crises, for example, in the 1970s and 1980s. It applies to studies of the role of socioeconomic understanding in establishing preconditions for institutionalized processes of compromising and consensus making in the postwar period in the Nordic and other small Western European countries. It also applies to policy studies of how environment, energy, and social and labor market policies are established and change within epistemic communities. Finally, it also applies to studies of European integration and international relations. In this way, both approaches have focused on the role of speech acts and texts in analyzing the articulation of politics; and while different concepts are put to use (ideas or discourse), both approaches take it for granted that language matters, just as it is accepted that politics takes place through the construction of interpretative frameworks. This finally places discourse analysis at the crossroads of theories of narration and political science as a theory of political communication. Within mainstream political science, there is a growing interest in the role of the genres (drama, history, and the epics) in producing political news or in how polities are framed and named. The use of narratives in political argumentation is explored in understanding news as a political institution and public policy as agendas for political discourse. Within political science, it is also becoming common to study the organizations in which ideas are produced (or consumed) and the epistemic communities through which ideas are debated and disseminated. The role of institutional entrepreneurs in universities, think tanks, expert systems, and the media, using narratives to frame and name political problems, is becoming an important topic.

Ove K. Pedersen
Copenhagen Business School
Frederiksberg, Denmark

See also Constructivism; Epistemological and Methodological Foundations; Institutional Theory; Institutions and Institutionalism; Policy, Discourse Models; Political Communication; Power

Further Readings

Campbell, J. L., & Pedersen, O. K. (in press). Knowledge regimes and comparative political economy. In D. Bèland & R. H. Cox (Eds.), *Ideas and politics in social science research*. Oxford, UK: Oxford University Press.

Edelman, M. (1988). *Constructing the political spectacle*. Chicago: University of Chicago Press.

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. London: Routledge.

Foucault, M. (1966). *Les Mots et les choses: Une archéologie des sciences humaines* [The order of things: An archaeology of the human sciences]. Paris: Gallimard.

Foucault, M. (1969). *L'archèologie du savoir* [The archaeology of knowledge]. Paris: Gallimard.

Hay, C. (2002). *Political analysis: A critical introduction*. Basingstoke, UK: Palgrave.

Kögler, H. H. (1999). *The power of dialogue: Critical hermeneutics after Gadamer and Foucault*. Cambridge: MIT Press.
Laclau, E., & Mouffe, C. (1985). *Hegemony and socialist strategy: Towards a radical democratic politics*. London: Verso.
Pedersen, O. K. (1991). Nine questions to a neo-institutional theory in political science. *Scandinavian Political Studies, 14*(2), 125–148.
Wetherell, M., Taylor, S., & Yates, S. J. (2001). *Discourse as data: A guide for analysis*. London: Sage.
Wetherell, M., Taylor, S., & Yates, S. J. (2001). *Discourse theory and practice: A reader*. London: Sage.

DISCRETION

Discretion is an issue of interest not only to political scientists but also to scholars of public administration, law, and organizations. In everyday language, using one's discretion refers simply to the employment of good judgment, caution, or care in decision making and, hence, is normally juxtaposed with acting according to formal standards or rules. Political scientists are interested in the exercise of discretion delegated to politicians by their electorate, but it is normally within the administrative realm that the issue of discretion (and the related matter of decision-making legitimacy) provokes most debate and discussion. This is because considerable discretionary authority tends to be granted to bureaucracies in order to make and implement policy, as legislatures are content to delegate the implementation of agreed policy to those at the "street level." For students of public administration, therefore, discretion can be defined as the autonomy extended to bureaucrats to implement or disregard a law in accordance with the principles established in a political arena for that law, based on values appropriate to the public interest. Those who exercise bureaucratic discretion range from effective standard setters (including regulators) who exercise considerable discretionary authority to those enjoying more limited authority when applying established standards to individual situations.

Historically, as a core concept in law and latterly in public administration, discretion has been examined by way of consideration of the different external mechanisms used for its limitation and control, including legislative oversight, executive command, and judicial review. As the scope of state power has expanded, it is the role of the judicial system in limiting discretionary behavior that, however, has generated an almost universal debate. Indeed, much of contemporary writing on the issue of discretion stems from legal theory with a strong focus on the role of law in policy implementation. For lawyers and legal philosophers, there is considerable discussion as to how and why the state's administrators, from police "on the beat" to judges in courtrooms, determine the public interest and exercise their discretion. This literature is concerned with the relationship between discretion and various conceptions of justice. Key questions include how much discretion, if at all, should be granted, how will it be granted, and, more fundamentally, who will decide on the public interest?

Arguments for and Against the Use of Discretion

The arguments for and against discretionary behavior in decision making are well developed. For advocates of discretion, its principal virtue is that it provides decision makers (politicians, judges, regulators, and civil servants) the ability to fill the gaps emerging between established rules of law. In other words, for a policy to be matched with the appropriate type of implementing mechanism, forms of discretion are required so that the spirit, if not the letter, of the law is adhered to. For this school of thought, discretion is an essential tool to help decision makers resolve problems in ways that can accommodate the parties involved and also soften the often harsh and indiscriminate impact of rules. They argue that rules limiting discretion may formally legitimate decisions, but such decisions may not be fair or may conflict with the public interest. Therefore, and in line with the theory of representative bureaucracy, administrators must have discretion as it allows them to produce results that reflect the public interest—that is, the values and beliefs held by the public they serve. It follows that officials should be provided with "formal" discretionary powers when options for policy implementation are written into a law. Thus, discretion forms an integral part of the legal process, which, among other things, allows for contexts and circumstances to be taken into account when assessing breaches of the law.

The arguments against discretion are also well rehearsed, particularly within legal philosophy where the rule of law constitutes a buffer between the citizen and the arbitrary exercise of power. In contrast with those who advocate the provision of discretion, advocates of rules and rule adherence (most prominently Kenneth Culp Davis) point to the fact that discretion facilitates an uneven or subjective implementation of the law by providing decision makers the ability to employ illegitimate considerations in the formulation of a decision. In such circumstances, decisions can be made using criteria that are in conflict with those established and envisaged by the rules, for example, nepotism or gender discrimination. Therefore, the enforcement of rules ensures that like cases are treated alike and avoids bringing a "private" dimension into decision making. Discretionary powers vary considerably and may in certain circumstances extend to rule breaking, arbitrary behavior, and injustices. It follows that the scope for inappropriate discretion (as opposed to necessary discretion as in the case of regulators) should be narrowed as much as possible by legislation.

The Nature of Discretion

Discretion is routinely contrasted with the rule of law—that is, those obligations and duties that constrain and determine the choice of action. There is much debate within public administration and legal philosophy as to whether discretionary behavior is rule-free or is in fact embedded within and influenced by rule frameworks (the famous "doughnut" conceptualization of discretion proposed by Ronald Dworkin in which discretion occupies the hole in the doughnut's center). Many scholars seek compromise and argue that dichotomies between rules and discretion fail to recognize the complex interdependence between the two concepts, particularly as a greater number of delegated bodies and actors with varieties of rule making and hence discretionary powers now exist in modern polities.

Within legal texts, there are two broad categories of discretion. The first is the *discretion to make rules*, which is normally the preserve of political institutions such as legislatures or local authorities (which are subject to electoral sanction) or indeed of other fora to which such functions are delegated. The second concerns the *discretion to find facts and interpret them as the law*—that is, the discretion to decide cases. This power is retained by courts and administrative agencies. Of course, the two are connected, and the central issue in the analysis of discretion is the extent to which legislatures, in the process of making law, constrain or provide discretion to courts and agencies to decide on cases. At its core, administrative law centers on the creation of rules that allow government and its agencies the freedom to achieve agreed goals while discouraging them from abusing this freedom.

Courts and judges tend to enjoy considerable discretionary powers in the enforcement and administration of the law. In this instance, discretion is most tangible when a judge, having consulted all relevant legal material, is left free by law to decide on a course of action. The type of public law used is also of relevance here, as it is generally thought that within common law legal systems, judges are afforded more discretion than their civil law counterparts.

Applications of Discretion

For scholars of political science and public administration, since the mid-1930s, research into administrative discretion has been primarily concerned with analysis of how the use of discretion by street-level bureaucrats leads to varieties of outcomes in public service delivery. Within the study of public management, a well-developed strand of research concerns the tensions inherent in the search for norms of value-free administration and the exercise of discretion in increasingly complex policy environments. More recently, the debate on discretion has become a core concern of regulation and regulatory governance, best captured by the literature on the "regulatory state" and its accountability. In particular, the work of independent regulatory agencies, which can be quasi judicial in character, raises important questions concerning the proper balance between formula-driven and discretionary approaches to regulation. Much recent scholarship has demonstrated a wide variety in the interpretation by regulatory agencies of their discretionary role and the principles on which their decisions are based.

The understanding and use of discretion depends on its context, and many tasks will inevitably remain too complicated to legislate for in a detailed manner, thus requiring the individual bureaucrat to respond on a case-by-case basis. It may also occur that bureaucrats are subject to contradictory rules and therefore have discretion to decide on which rules to implement. Discretion also tends to be of more significance in contexts where resources are in short supply but demands are unlimited, providing opportunities for discrimination as well as rationing. By way of contrast, contracts between public authorities and private companies over service delivery (such as waste disposal or road building) seek to ensure that discretion rests solely with the principal rather than with the agent. Discretion sharing also occurs in collaborative forms of organizations, such as networks.

As street-level bureaucrats often provide the public face of government, their activities have a symbolic import that stretches beyond the service in question. Their decisions can reinforce or undermine political objectives. For political and administrative managers, therefore, problems of asymmetrical information mean that controlling the discretionary activities of street-level bureaucrats through normal superior–subordinate routines is problematic. This does not stop street-level bureaucrats from being subjected to ever-increasing amounts of rules. Therefore, apart from the external means of restricting administration discretion (above), internal means such as administrative ethics, peer review, and an emphasis on professional norms are emphasized. The development of common public service values through induction and training is also routinely identified as a means of ensuring that discretionary behavior operates within commonly shared norms and understandings.

Organizational theory holds that wherever work is delegated, there is some loss of control by the delegating person or institution. Also, the greater the degree of task complexity, the greater the need for discretion by those involved. Thus, discretion is closely connected to levels of trust within organizations, with managers seeking a virtuous circle of high trust and high discretion in performance. One strand of research on organizations is increasingly interested in the evolution of rules and discretion within hierarchical organizations. They examine how organizations may shape discretion by imposing constraints on the way in which members use it. The formal and informal means through which discretionary behavior is tolerated or sanctioned is also an emerging field of inquiry.

Within organizational theory, it is also proposed that culture plays a role in the use of discretion, and work from this field of inquiry has identified how increased bureaucratic formalization tends to diminish discretion. Organizational sociologists are also keenly interested in the concept, proposing that a greater understanding of the determinants of discretionary behavior requires analysis of the social norms prevalent within an organization's operating environment. The dilemmas inherent in the exercise of discretion thus help explain the origins of trust, behavioral norms, and boundaries of autonomy.

Muiris MacCarthaigh
Institute of Public Administration
Dublin, Ireland

See also Bureaucracy; Bureaucracy, Street-Level; Organization Theory; Regulation; Rule of Law

Further Readings

Davis, K. C. (1969). *Discretionary justice: A preliminary inquiry.* Baton Rouge: Louisiana State University Press.

Dworkin, R. (1977). *Taking rights seriously.* Cambridge, MA: Harvard University Press.

Hawkins, K. (Ed.). (1992). *The uses of discretion.* Oxford, UK: Clarendon Press.

Herring, E. P. (1936). *Public administration and the public interest.* New York: McGraw-Hill.

Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public service.* New York: Russell Sage Foundation.

Discriminant Analysis/ Classification

See Categorical Response Data

DISCRIMINATION

Originating from the Latin word *discriminare* (distinguish, separate), in contemporary societies, the term *discrimination* is applied to related but varying phenomena about the content of which there is no general agreement. The least controversial meaning is the intended or accomplished differential treatment of persons or social groups for reasons of certain generalized traits. The targets are often minorities, but it may also be majorities, as Black people were under apartheid in South Africa. For the most part, discrimination implicitly entails the idea of acts resulting in a disadvantage for the target of such action. More precise terms are *intentional* or *willful discrimination.* An ever-growing number of terms has been coined to label phenomena of this kind related to groups, traits, or properties, such as race (racism), religion (e.g., anti-Semitism, Islamophobia), sex (sexism), sexual orientation (e.g., homophobia), class (classism), age (ageism), disability (ableism), a person's stature (heightism, discrimination against short people), physical appearance (lookism), overweight (weightism), and many more. Although there are many other forms of discrimination, not all of them have been identified as specific targets of social action. In the following sections, the different elements of such a definition, their controversial treatment in the social sciences, and their political implications are discussed.

Controversies in Defining Discrimination

While the definition of discrimination as differential treatment on the basis of generalized traits is relatively straightforward, a number of additional elements that might be added to the definition are controversial.

Must Behavior Be Unjustified?

As Michael Banton (1994) notes, "One analytical tradition excludes considerations of immorality and illegality altogether for the sake of conceptual clarity and to avoid contradicting opinions about their justification" (p. 1). But politicians argue that the definition should serve the purpose of identifying a social problem and thus restrict the term to *unjustified* behavior.

Are There Legitimate Exceptions?

Should references to *legitimate exceptions* be included? Many scientific and political definitions of discrimination contain conditions under which differential treatment is deemed undue versus those under which it is deemed admissible. Discrimination here relates only to unequal treatment that is based on criteria lacking objective or rational justification. Aaron Antonovsky (1960), for example, speaks of "effective injurious treatment of persons on grounds rationally irrelevant to the situation" (p. 81). The weakness of this definition lies in its assumption that rationality in the Weberian sense can define uniform collective goals that should override the legitimacy of egotistic, individual claims that run counter to them. Conflicts over protectionist discrimination against foreign companies illustrate that competitors, consumers, state economic planners, environmentalists, and Third World and workers' rights activists may have conflicting particular interests precluding a single universal rationality. What can be said is that discrimination is seen to exist when reasons for actions are perceived as illegitimate by relevant groups. But concepts of legitimacy vary between and within societies and change over time. Barring women from the ballot was long considered to be justified, and even in Western democracies, full formal economic rights for women were established only in the course of the 20th century. In feudal societies, it was generally accepted that access to positions of leadership was made difficult for, or denied to, persons of lowly origin, regardless of their ability or merit. Only in the modern age did the bourgeois classes question the nobility's inherited privileges. Since modern societies consider themselves universalistic in the sense that only merit shall determine social positions, a convergence of public opinion has occurred to exclude ascriptive (hereditary, unchosen, and unchangeable) criteria in favor of achievement only.

Human Rights or Positive Law?

Does any contravention of the universal principle of equal rights constitute discrimination regardless of the traits involved or does the definition relate to an *inventory of specific* qualities? Brandishing discrimination, activists often invoke natural law or general fundamental human rights.

Wary of potential liabilities resulting from open definitions, legislators prefer to enact positive law that defines specific new rights rather than relying on concepts of natural law.

Are There Exemptions for Private Actors?

Which *classes of actors* fall under the scope of the concept? While there is broad agreement that governments and para-statal agencies should not discriminate, private sector economic actors often successfully seek exemption on the ground that nondiscrimination will disadvantage them (e.g., by resulting in lower productivity, by overriding customers' preferences). In general, definitions are ambiguous as to their application in the realm of private life in which differential treatment based on personal preferences is widespread but considered to be legitimate.

Is Discrimination an Attitude or an Act?

Does a mere *attitude*, such as the intention to treat someone differentially, constitute discrimination, or must a tangible disadvantage following *accomplished action* occur? Much research in social psychology focuses on the conditions under which members of one group develop unfavorable attitudes toward members of other groups, a process labeled outgroup discrimination. Other disciplines such as the economic, political, and social sciences have traditions of observing discriminatory acts and their detrimental consequences. Discriminatory attitudes do not always lead to corresponding acts.

The Contingency of Discrimination

Discrimination is a *contingent* phenomenon since acceptance of universalistic principles and awareness of inequality have developed historically. Statistical evidence can demonstrate certain forms of social inequality while explanations remain contested. Demands that discrimination be recognized are made in contexts of power relations and often succeed only as a result of pressure groups claiming minority rights against vested interests. Discrimination thus is as much a state of political consciousness as it is an ontological matter. This holds all the more true at the level of legal codification. Specific legislation has been enacted in some countries, for example, with regard to mandating equal treatment for women, gays, and people with disabilities; however, differential treatment on grounds of foreign citizenship is considered defensible for raisons d'état and has failed to achieve recognition as discrimination. A glance at the fate of left-handed people, vegans, and others reveals that traits and categories exist that are connected to some kinds of inequality but that have not yet been associated with discrimination.

Types of Discrimination

Statistical discrimination is considered to occur when distinctions are made between demographic groups based on aggregate characteristics rather than on an actor's individual preference and prejudice. While a group is known to possess certain characteristics, an actor unaware of the personal characteristics of an individual treats him or her as an average member of a group. But the perception of similar phenomena of this type is inconsistent. Ethnic profiling—that is, differential treatment of an ethnic or racial group by security personnel, may be considered discriminatory even if evidence existed to support it as a rational strategy of law enforcement, because members of certain groups figure high in criminal statistics. When employers promote equally qualified women to leading positions less frequently because experience shows a higher likelihood of interruption of employment due to pregnancy or child care, political reactions often are ambiguous: This can be seen as illegitimate with regard to the individual person, whereas employers defend it as rational behavior maximizing productivity. By contrast, it is less noncontroversial that insurance companies grant preferential conditions according to age, sex, occupation, or other criteria based on evaluations of risk statistics.

While *intentional discrimination* is located at the level of individuals, *institutional discrimination* denotes explicit policies of social institutions that exclude, impede, or otherwise harm certain groups irrespective of adverse attitudes of implementing agents. Famous examples are laws restricting rights of racial or ethnic minorities or banning women from politics. Recently, courts in Europe have held proceedings against laws barring foreigners from particular professions. By contrast, *structural* (i.e., indirect) *discrimination* characterizes policies that are neutral in intent and in implementation but still result in adverse effects for minorities. In the literal sense, structural discrimination constitutes the very opposite of discrimination and it is, therefore, most controversial. Its proponents hold that states based on universal principles have to bring about equal life chances for all, which includes the duty on the part of institutions, for example, in education, to compensate proactively for the inequalities minorities have experienced from discrimination suffered in the past or in other social systems. They argue that the failure to do so constitutes discrimination.

Explanations of Discrimination

Social psychological explanations of discrimination based on social identity theory presume that humans rely on the groups they belong to for a part of their identity. Belonging to a group that is more prestigious and powerful than others boosts one's sense of self-esteem. Discrimination in the sense of debasing and impairing outgroup members or declining them access to resources and wealth serves the purpose of strengthening the relative position of one's in group and also indirectly boosts individual self-esteem. Empirical studies confirm that persons with a low sense of social recognition display more outgroup devaluation and *group-focused enmity* based on an ideology of human inequality (Wilhelm Heitmeyer, 2002). Negative attitudes toward different outgroups (ethnic and religious minorities, women, and people who are disabled or homeless) strongly intercorrelate, indicating the unspecific nature of discrimination. Social dominance theory argues that societies as a consequence of evolution generally develop hierarchical structures and humans have a disposition of dominance orientation and willingness to discriminate, but, in addition, a high group status exerts a reinforcing influence on this attitude.

Beyond statistical discrimination, *economic explanations* argue that a part of the difference between groups in income, prestige, and status achievement is due to differences in productivity. Women, for example, are assumed to choose occupations compatible with family demands and akin to female role models that are on average paid worse and to invest less in training since their returns would be smaller than men's. Differential treatment based on productivity gradients is held to be legitimate. According to Gary Becker (1971), illegitimate differences are explained by "tastes of discrimination" on the part of employers, customers, and landlords. Since such tastes imply inefficient attribution of factors of production, they cause disadvantages on perfect markets and are therefore deemed to vanish over time. It has been shown, however, that in majority–minority situations, under state control, and in other market irregularities discrimination due to taste may persist. Economic explanations have been criticized for failing to address the root causes of differential preferences along gender, class, and ethnic lines and the mechanisms behind discrimination.

Sociological analyses have shown widespread ideological constructions of superiority of certain groups, for example, "gender status beliefs," legitimizing unequal access to status positions. One area of research is dedicated to the intergenerational stability of class positions. Pierre Bourdieu (1980), for example, argues that powerful groups define a specific taste that, internalized at an early age, serves as a lifelong means of distinction between classes and regulates status assignment in contradiction to meritocratic principles of modernity.

Wilhelm Heitmeyer and Kurt Salentin
Bielefeld University
Bielefeld, Germany

See also Equality; Gender; Social Exclusion

Further Readings

Antonovsky, A. (1960). The social meaning of discrimination. *Phylon, 21*(1), 81–95.

Banton, M. (1994). *Discrimination.* Buckingham, UK: Open University Press.

Becker, G. S. (1971). *The economics of discrimination.* Chicago: University of Chicago Press.

Bourdieu, P. (1980). *La distinction: Critique sociale du jugement* [Distinction: A social critique of the judgement of taste]. Paris: Minuit.

Heitmeyer, W. (Ed.). (2002). *Deutsche Zustände* [German state of affairs]. Frankfurt, Germany: Suhrkamp.

Pincus, F. L. (1999). From individual to structural discrimination. In F. L. Pincus & H. J. Ehrlich (Eds.), *Race and ethnic conflict: Contending views on prejudice, discrimination, and ethnoviolence* (pp. 120–124). Boulder, CO: Westview.

Discursive Institutionalism

The term *discursive institutionalism* is an umbrella concept for the wide range of works in political science that focus on the substantive content of ideas and the interactive processes of discourse by which they are generated and communicated in given institutional contexts. This makes for a definition that is somewhat broader than when the same or similar terms are used by those who focus primarily on the substance of political agents' ideas in the ideational turn or on the text in discourse analysis. Below, relations to other institutionalist approaches, varieties of this concept, and their applications in political science are discussed.

Discursive institutionalism can be considered the fourth "new institutionalism," because it is a distinctive approach that contributes to our understanding of political action in ways that the other three neo-institutionalisms—rational choice, historical, and sociological—cannot. It provides insight into areas of politics that political scientists working within the other neo-institutionalist traditions have long neglected: the role of ideas in constituting politics and in reconstructing political interests and values, the power of discourse for persuasion in political debate and legitimation in democracy, the importance of ideas and discourse in explaining the microdynamics of change in rationalist interests and structures, historical paths and regularities, and cultural frames and norms.

The word *institutionalism* in the term, moreover, suggests that this approach is not only about the discursive construction and communication of ideas but equally about the institutional context in which and through which ideas are communicated via discourse. This institutional context refers first and foremost to the structure, construction, and communication of meaning. But it can also be understood as the background information provided by the other three neo-institutionalisms in political science with which discursive institutionalists may engage and from which they often emerge.

Varieties of Approaches

Discursive institutionalism encompasses a wide range of approaches across fields of political science, from constructivism in international relations to the ideational turn in American political development and comparative politics, from the deliberative democracy of political theory to the analysis of campaign rhetoric and media discourse in American electoral politics, and from the discourse analysis of critical legal studies or poststructuralist theory to that of feminist gender studies. In such approaches, when discursive institutionalists focus on ideas, they may consider different types of ideas—cognitive ideas justified in terms of interest-based logics and/or normative ideas legitimated through value-based logics of appropriateness—at different levels of generality, from policy ideas to programmatic ideas or from paradigms to deeper philosophical ideas. When they are more concerned with discourse, they may concentrate on the representation of ideas through frames, narratives, myths, collective memories, stories, scripts, and the like or on the discursive processes by which such ideas are constructed and communicated in the public sphere.

Discursive institutionalists may also be divided over which part of the public sphere they ought to investigate. Some are more interested in the *coordinative* discourse of the policy sphere, in which policy actors construct ideas through epistemic communities of loosely connected individuals with shared ideas about a common policy enterprise, advocacy coalitions of more closely connected elites with shared ideas and access to policy making in local and national politics, advocacy networks of activists contesting ideas in international politics, and entrepreneurs or mediators who serve as catalysts for change as they articulate the ideas of discursive communities. Other discursive institutionalists are more focused on the *communicative* discourse of the political sphere, in which political actors discuss, deliberate, and legitimate the ideas developed in the coordinative discourse. These actors include elected officials, opposition parties, or spin doctors, civil servants, and/or policy actors involved in the mass electoral process of public persuasion. But they also involve opinion leaders, public intellectuals, and the media engaged in public debates; organized interests in specialized policy forums; members of civil society engaged in grassroots organizations and social movements; or mini publics participating in citizen juries and issue forums, as well as citizens whose voices are heard not only in opinion polls but also in votes—where actions speak even louder than words. The communicative discourse itself may be seen as top-down, as elites seek to form mass public opinion or legitimize public policy; as bottom-up, as social activists attempt to influence national and international debates; at the top among policy and political actors, as elites debate ideas in national or global policy forums; or at the bottom, among everyday actors engaged in deliberative democracy or in expressing ideas through their everyday practices.

Discursive Institutionalism in Institutionalist Perspective

All discursive institutionalists define institutional context first of all in terms of the *meaning context* in which ideas and discourse make sense, such that speakers "get it right" in terms of the ideational rules or rationality of a given setting by addressing their remarks to the "right" audiences, at the "right" times, and in the "right" ways. But beyond this, they may differ in terms of how they refine their definitions of institutional context and their objects of explanation, depending on the background information they add from one or another of the other neo-institutionalist traditions.

For example, discursive institutionalists who engage with rational choice institutionalism also often use the language of interest maximization and incentive structures. But unlike rational choice institutionalists, who tend to put interests first, leaving aside ideas as little more than mechanisms to choose among objective interests or as focal points to choose among equilibria, discursive institutionalists put ideas first, as constituting or constructing subjective interests that encompass a wider range of strategic ideas and social norms than most rational choice institutionalists admit. As for institutions, rather than the rational choice institutionalists' "neutral" incentive structures, discursive institutionalists see them as infused with ideas and values or collective memories that make them objects of trust or mistrust and changeable over time as actors' ideas and discourse about them change along with their performance. Moreover, instead of attributing interests to "rational" actors who are assumed to live in a world of knowable risk, discursive institutionalists investigate the actual ideas about interests of real actors who live in a world that is often characterized by radical or immeasurable uncertainty. Finally, against the rational choice institutionalists' view of talk as by definition "cheap" and actions as "speaking more loudly than words," discursive institutionalists take what people say about their actions as well as the exchange of ideas in public debates and discourse as key to explaining politics, as the basis of leadership through the top-down processes of mass persuasion and of democracy through the bottom-up processes of opinion formation and deliberation.

Discursive institutionalists who engage with sociological institutionalism have much less to disagree with since sociological institutionalists already use the language of ideas, norms, and discourse. Moreover, for the most part, they share a constructivist ontology about the nature of interests and institutions. For both, in place of rationalists' assumptions about material interests as determinant factors in individual action, they consider political agents' ideas about interests as well as norms in response to material reality. For discursive institutionalists, the main sources of disagreement arise with those sociological institutionalists who come down more on the side of institutional stasis, by emphasizing how cultural frames constrain ideas, discourse, and action, as opposed to the institutional dynamics that comes from the reframing of action through ideas and discourse.

Discursive institutionalists who engage with historical institutionalism have similar concerns about the static nature of approaches that emphasize institutional path dependence and see change as the result of unexplained exogenous shocks. They find, moreover, that even though the recent shift in historical institutionalism from exogenous to incremental or evolutionary theories has gone a long way toward endogenizing change, this only does more to describe *what* changes occur than to explain *why* change occurs. Any such explanation requires a move from a focus on structure to agency, to explore the microfoundations of the macropatterns. But instead of continuing to turn for such agency either to the incentive-driven rational agents of rational choice institutionalism or to the culturally framed social agents of sociological institutionalism, discursive institutionalists argue that historical institutionalists would do better to look to the ideas and discourse of sentient—or reflexive, meaning self-reflective thinking and speaking—agents.

For discursive institutionalists, the problem with the three other neo-institutionalisms is that the structures—of incentives, culture, or path dependence—that make up the institutional context act mainly as external constraints on action. Instead, discursive institutionalists treat institutions at one and the same time as given (i.e., as structures that are the context within which agents think, speak, and act) and as contingent (i.e., as constructs that result from agents' thoughts, words, and actions. Here, continuity and change in institutions can be explained by the dynamic processes through which agents use not only their *background ideational abilities*—a generic term for John Searle's *background abilities* or Pierre Bourdieu's *habitus*—to create and maintain not only their institutions but also their *foreground discursive abilities*—akin to Jürgen Habermas's *communicative action*—to communicate and deliberate about taking action collectively to change (or maintain) those institutions.

The relationship between discursive institutionalism and the other neo-institutionalisms is complicated, therefore, since discursive institutionalists may accept some of the language and logic of the other approaches even as they critique them. With regard to rational choice institutionalism, discursive institutionalists sometimes take its account of interest-based ideas as a shortcut to that which can be expected (albeit not predicted) although even the expected is always open to investigation. With regard to sociological institutionalism, discursive institutionalists often take its account of culture, norms, and frames as a useful starting point for a closer examination of the dynamics of change (or continuity). With regard to historical institutionalism, discursive institutionalists help explain the dynamics of change in historical structures through an agent-centered, meaning-based logic of communication. By the same token, however, historical institutionalism can lend a hand to discursive institutionalism, by helping show how formal institutions may shape discursive interactions. For example, compound polities, in which governing activities are dispersed among multiple authorities, tend to produce more elaborate coordinative discourses among policy actors, whereas simple polities, in which governing activities tend to be channeled through a single authority, facilitate more elaborate communicative discourses to the public.

Discursive Institutionalist Methods

But how can discursive institutionalists demonstrate that ideas and discourse make a difference for political action, let alone have a causal influence? Political scientists often use methods based on comparative case studies and process tracing to demonstrate how ideas and discourse are tied to action. This may involve pointing to the ideas that serve as before-the-fact guides for action, as after-the-fact sources of justification and legitimation for action, or as ongoing frames through which actors understand as well as explain their actions. But it may instead involve considering the discursive exchange of ideas, as actors in debates or negotiations change their ideas as a result of the process of persuasion—something that Habermas calls "arguing"—in contrast to "bargaining." Another such method involves showing that ideas and discourse can be independent variables by demonstrating that no other neo-institutionalist explanation, whether in terms of rationalist interests, historical path dependencies, or cultural norms, does a better job of accounting for the clear changes (or continuities) in political action. Examining matched pairs of cases in which everything is controlled for except the ideas and discourse can also usefully show when these phenomena make a difference. Finally, although qualitative methods are most prevalent, discursive institutionalists also employ a variety of quantitative techniques such as surveys, opinion polls, and fuzzy sets to establish the influence of ideas and discourse.

Vivien A. Schmidt
Boston University
Boston, Massachusetts, United States

See also Constructivism; Discourse Analysis; Historical Memory; Institutions and Institutionalism; Path Dependence; Policy Framing; Rational Choice

Further Readings

Blyth, M. (2002). A theory of institutional change. In M. Blyth (Ed.), *Great transformations* (pp. 17–45). New York: Cambridge University Press.

Campbell, J. L. (2004). The problem of ideas. In J. L. Campbell (Ed.), *Institutional change and globalization* (pp. 90–123). Princeton, NJ: Princeton University Press.

Campbell, J. L., & Pedersen, O. (2001). Introduction. In J. L. Campbell & O. Pedersen (Eds.), *The rise of neoliberalism and institutional analysis* (pp. 1–24). Princeton, NJ: Princeton University Press.

Rothstein, B. (2005). Trust and collective memories. In B. Rothstein (Ed.), *Social traps and the problem of trust* (pp. 148–166). Cambridge, UK: Cambridge University Press.

Schmidt, V. A. (2006). Theorizing democracy in Europe. In V. A. Schmidt (Ed.), *Democracy in Europe: The EU and national polities* (pp. 219–266). Oxford, UK: Oxford University Press.

Schmidt, V. A. (2008). Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science, 11,* 303–326.

Seabrooke, L. (2006). *The social sources of financial power* (chap. 2). Ithaca, NY: Cornell University Press.

Discursive Policy Analysis

Discursive policy analysis focuses on discussion or debate as the main object of scrutiny. Originally, the word *discourse* was derived from the Latin noun *discursus*, which means "running to and fro." It is this dynamic aspect that we still find in discursive policy analysis, which often concerns itself with the development of meaning. How do specific meanings evolve during the policy process? Why do certain terms pop up more often? Can we see how coalitions are organized around these terms? Why does a policy process move first this way, then that way? This entry first gives a short overview of the various meanings associated with the term *discursive policy analysis*, then elaborates on the relationship between discourse and practice and language and power, and ends with a short note on the normative or prescriptive potential of discursive policy analysis.

The Meaning of Discursive Policy Analysis

The discursive analyst takes discussion as the object of analysis and then sets out to trace linguistic regularities within those debates. These linguistic regularities are generally called discourses. However, the term *discourse* is used in different ways, ranging from the speech–act theory to the "archeology" of Michel Foucault.

What the approaches share is an anti-essentialist ontology. Discursive analysis can be placed within the constructivist or interpretive tradition of the social sciences: It assumes that language is constitutive of social realities and takes a critical stance toward a singular notion of an objective "truth" waiting to be discovered. This is not to assume that phenomena exist only in language (as critics tend to state), but it does suppose that we can perceive phenomena only by linguistically making sense of them. This assumption of multiple social realities is of particular importance for the policy analyst, because, in the words of Carol Bacchi (2000),

> The premise behind a policy-as-discourse approach is that it is inappropriate to see governments as responding to "problems" that exist "out there" in the community. Rather "problems" are "created" or "given shape" in the very policy proposals that are offered as "responses." (p. 48)

Societal events are almost by definition ambiguous and so complex that they can never be fully described. Every description or narrative gives a particular perspective on the problem; indeed, the very perspective will determine whether a particular phenomenon will be seen at all as a problem that requires political action. Words, catchphrases, metaphors, and story lines frame societal problems, and as this framing influences our answer to the question, "Who gets (and who should get) what, when, and how?" it is inherently political.

As language is not transparent—we cannot "see through" it—there is a continuous interaction between the terms in which a policy is described and the way we think about it. In their famous book *Metaphors We Live By* (1980), George Lakoff and Mark Johnson explained how describing something in terms of something else (what they call the "thisness of a that") structures our thought processes. For instance, it matters whether we describe integration in terms of a melting pot, an invasion, or even a war, as the description will shape not only our understanding of what it is but also the ways we respond to it. Moreover, a particular policy will always come with categories and target groups and descriptions of deserving and undeserving people, thereby creating its own political reality. In the words of Norman Fairclough and Ruth Wodak (1997),

> discourse is socially constitutive as well as socially shaped: it constitutes situations, objects of knowledge, and the social identities of and relationships between people. It is constitutive both in the sense that it helps to sustain and reproduce the status quo, and in the sense that it contributes to transforming it. (p. 258)

The central question then becomes how certain meanings are created, challenged, or sustained.

Language in Use

To answer that question, studying language alone will often prove insufficient. After all, language does not "float" in society but is uttered in particular contexts that influence its meaning; the words "Yes, I do" have an entirely different meaning at the altar in a church than at the card table in a pub. Likewise, a particular policy might be documented in papers and reports, but in most cases, its meaning will be formed and solidified during implementation, in the interaction between street-level bureaucrats and (would-be) target groups. Meanings are not only localized in language but also in symbolic artifacts and the physical environment in which the interaction takes place. Discourse analysis, then, is contextual; it is the study of language in use.

Such a statement could be placed in a Foucaultian tradition of discursive analysis that tried to "overcome the traditional distinction between what one *says* (language) and what one *does* (practice)" (Stuart Hall, 2001, p. 72). Thus, Maarten Hajer (1995) defined discourse as a set of ideas, concepts, and categories through which meaning is given to social and physical phenomena. That "ensemble" is produced and reproduced through an identifiable set of "practices." There is, however, a second problem of definition, because *practice* is a catch-all term as well. According to Etienne Wenger (1998), *practice* is "doing in a historical and social context that gives structure and meaning to what we do" and includes not only "the language, tools, documents, images, symbols, well-defined roles, specified criteria, codified procedures, regulations and contracts" but also "the implicit relations, tacit conventions, subtle cues, untold rules of thumb, recognizable intuitions, specific perception, well-tuned sensitivities, embodied understandings, underlying assumptions, and shared world-views" (p. 47).

This leads to the paradoxical situation that discourse can include practice, that practice also includes discourse, and that both terms can be used to describe almost anything. It seems justified to ask, paraphrasing Aaron Wildavsky, "If discourse is everything, maybe it is nothing?"

Words and Interests

When reading the literature on discursive policy analysis with its contradictory and often implicit definitions, one might be tempted to answer affirmatively. By making the explanandum (social reality) into an explanans (practice), a structure–agency dilemma can be easily kept out of sight. No wonder then that Bacchi (2000) argues that discursive analysts tend to be positioned on the left of the political spectrum, to come in with an agenda for change and to define discourse in a way suiting their political purpose as constraints on change.

While the truth of this observation can hardly be denied, the ability to answer not only "what" but also "why" and "how" questions remains a crucial strength of discursive policy analysis, but this requires explicit attention for the way in which language is related to power and vice versa. To what extent are subjects "controlled" by the unconscious use of language in context that in itself incorporates particular power differences or, on the contrary, to what extent do subjects control their environment by strategically using discourse? And if subjects are captivated in the worldview of dominant discourses, how would transformation be possible? Again, various discursive policy analysts would answer these questions in different ways; whereas some scholars (e.g., the political scientist Murray Edelman) emphasize the strategic and conscious use of symbolic language by those in power, practice researchers point to the way in which the relationship between actors is constituted by the contextualized practices in which they act.

Both practices and discourses have meaning only to the extent that they are taken up by the actors involved in a particular process. Whereas subjects will almost always draw on stock texts and practices and on memory, conventions, and historical references, situations might arise in which these default narratives are no longer sufficient and can be challenged by new discourse coalitions. During these moments of dislocation, discursive policy analysis can be helpful not only for studying how power is given shape in particular practices but also to understand how interests come about in the first place. Whereas money and power are constant factors in politics, as both means and ends, it is only via language that actors can set their priorities and determine what their interests are. Vice versa are the chances for a particular discourse to become dominant, or at least partly dependent, on unevenly distributed means for symbolic distribution (e.g., easy access to the media).

What About the "Real" Problems?

The adage of "speaking truth to power" seems hard to combine with the ontological assumptions of discursive policy analysis, but this does not mean that all normative or prescriptive ambitions should be set aside. Various analysts have asked themselves whether discursive interpellations could also be actively promoted by active discourse management. Policy problems are not discovered "out there," but they are linguistically constructed. Yet some of these linguistic constructs may be more helpful for understanding complex phenomena and thus for developing effective policy practices. Whereas this is obviously a complicated endeavor, it also raises questions about the role of the policy analyst. Rather than regarding oneself as an objective provider of knowledge, the policy analyst becomes an active mediator in the complex interaction between politics, science, and societal groups. The art of discursive policy analysis could then be considered as, in the words of Piet Hein, "the solving of problems that cannot be expressed until they are solved."

Maarten Hajer and Wytske Versteeg
Universiteit van Amsterdam
Amsterdam, Netherlands

See also Analytic Narratives: Applications; Data, Textual; Discourse Analysis; Discursive Institutionalism; Identity, Social and Political; Interviewing

Further Readings

Bacchi, C. (2000). Policy as discourse: What does it mean? Where does it get us? *Discourse: Studies in the Cultural Politics of Education, 21*(1), 45–57.

Fairclough, N., & Wodak, R. (1997). Critical discourse analysis. In T. A. van Dijk (Ed.), *Discourse as structure and process: Discourse studies, a multidisciplinary introduction* (pp. 258–284). London: Sage.

Hajer, M. A. (1995). *The politics of environmental discourse: Ecological modernization and the policy process.* Oxford, UK: University of Oxford Press.

Hall, S. (2001). Foucault: Power, knowledge and discourse. In M. Wetherell, S. Taylor, & S. J. Yates (Eds.), *Discourse theory and practice: A reader* (pp. 72–81). London: Sage.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge, UK: Cambridge University Press.

DISSATISFACTION, POLITICAL

Political dissatisfaction is the attitudinal expression of unhappiness or lack of satisfaction based on the belief that the government or incumbent policies and their outcomes are falling short of the citizen's preferences or expectations, resulting in a negative feeling toward the government and the incumbents. This definition implies the following from a theoretical point of view: First, it means that this is an attitude related mostly with the evaluation of policies and their outcomes. Second, it implies that citizens are able to distinguish and identify those policies and their consequences, evaluate them, and compare them with their given preferences. Third, the definition also signifies that citizens attribute these policy decisions to specific political authorities who are responsible for selecting and implementing them. Fourth, it means that citizens think that there are alternative policies that better suit their problems and expectations, opening the possibility of giving the support to other contenders. Below, the origins and various interpretations of this concept, its relationship with similar ones, and possibilities of its operationalization are discussed.

Origin and Historical Evolution of the Concept

The political dissatisfaction concept is linked theoretically to David Easton's model of system support. According to Easton, it is citizens' evaluation of political outcomes that determines the degree to which public authorities, governments, and political systems as a whole are supported by mass publics. Easton considers that political dissatisfaction is part of the political support, and, more concretely, part of the "specific support" (or the lack thereof) that lies in the relationship with the satisfaction that members of a system feel they obtain from the perceived outputs and performance of the political authorities if they are opposed to them, disquieted by these policies, and dissatisfied with their conditions of life.

This contribution has led many scholars to misinterpret the concept that remained conceptually and methodologically undefined. This is why there is a diverse use of the concept and many variations of its operationalization in the literature, leading to important theoretical and empirical confusions. Often, this problem arises when linking political dissatisfaction with political mistrust or cynicism toward the entire political system, making all these attitudes part of the same attitudinal continuum pertaining to the same one-dimensional concept of political support. A good example of this occurred when William Gamson provided a general theoretical notion of the process underlying the development of political distrust and political disaffection. He maintained that dissatisfaction can lead to a more general problem of institutional and system distrust when an undesirable outcome is seen as producing a general distrust vis-à-vis political authorities and the whole representation process. This interpretation also can be found in the writings by Robert Dahl and, more recently, Leonardo Morlino, when they assert that political dissatisfaction results from the perception of government's general inability to deal effectively with problems regarded by citizens as important.

However, these scholars and other authors following this line of thought forget that each political object, including political authorities, can be the object of diffuse or specific support. What distinguishes the type of support is not the object of the support but the nature of the support—that is, if it is conditional (instrumental) or unconditional. Specific support can be object specific in two ways: First, people are assumed to be aware of the political authorities responsible for decision making; second, it takes into account the perceived decisions, policies, actions, proclamations, and style of the authorities. In this sense, the members of a political system can perform a rational calculation to find out whether the authorities' actions address their needs and demands. Under such conditions, specific support, which can only exist in societies whose institutions allow authorities to be held accountable for their actions and the resulting policies, will fluctuate according to people's perceived benefits and preferences. So dissatisfaction with specific policies and the authorities approving and implementing them, which is the result of instrumental calculations conditioned by people's preferences, does not necessarily affect the diffuse support accorded to the governmental institution as such. Furthermore, this relationship between output and general political trust is empirically questionable. Recent literature has demonstrated why this may not happen in many cases (see, e.g., Russel J. Dalton, 2004; Richard P. Gunther & José Ramón Montero, 2006). Finally, Gamson's developmental model does not incorporate a concern with cause and effect or the intricate set of interactions that may occur over time between policy dissatisfaction, distrust, and social or historical events. Declining levels of political trust may be the result of increased policy dissatisfaction, or it may be the other way around, or there might be an independent factor responsible for the decline in both of these variables. Some societal factor may cause an increase in political cynicism, which in turn acts as an intervening variable to bring about a rise in dissatisfaction with government policies. Alternatively, cynicism and dissatisfaction could both be connected with some changing social conditions and, therefore, be only spuriously related to each other.

Some authors, however, depart from this misleading interpretation of the concept. In this line of thought, Guiseppe Di Palma developed at the beginning of the 1970s a definition of political dissatisfaction very close to the one presented above. According to Di Palma, political dissatisfaction is an expression of displeasure resulting from the belief that the performance of the government or political system is falling short of citizens' wishes or expectations. This is an important clarification because of the following two reasons. First, it points out that political dissatisfaction is the product of the difference between expectations and policy outcomes, and, as is well-known by now, expectations are endogenous to one's own political system and the fulfillment of demands provided by the incumbent government; second, it focuses on the policies produced by a given government or incumbent actors, implying some sort of short-time evaluation. Still more important is the contribution by Arthur Miller, one of the protagonists of the well-known Miller–Citrin controversy that took place during the mid-1970s. Miller, distancing himself from Jack Citrin's (1974) comment, maintains that satisfaction with the political authorities and political trust are two theoretically different things that should be separated in empirical research. The contribution by Miller (1974) has become crucial over the years. He is signaling that the applicability of the concept of political support depends on the validity of the assumption that people can be aware, however vaguely, of a relationship between their needs, wants, and demands, on the one hand, and the actions of the political authorities on the other. The relationship has to be such that the citizens perceive, whether correctly in some objective sense or not, that the fulfillment of their needs and demands can be associated with the authorities in some way. Finally, this way of understanding and defining political support and political dissatisfaction has been continued by Barbara Farah, Samuel Barnes, and Felix Heunks (1979) when they affirm that the concept of political support is based on the belief that the political order should be responsive to citizens' demands and that these demands are expected to result in governmental outputs. It is the responsibility of the elected officials to translate these demands into public policy, and subsequently, their performance can be evaluated by the general public. Dissatisfaction, then, emerges when performance is viewed negatively. This is the line of thought of more recent works by Mariano Torcal and José Ramón Montero (2006), although they make the concept of political dissatisfaction interchangeable with the concept of political discontent.

The final point important to mention here is that this attitude is the result of evaluations of incumbents and policies. It tends to produce a critical support for concrete authorities, being strongly related to the electoral support of these authorities. So the most concrete and conspicuous consequence of political dissatisfaction is electoral change, either approving or choosing other parties or producing electoral abstention. This is the most frequent result of political dissatisfaction. Another question, even though not fully investigated, concerns the potential consequences of political dissatisfaction on other attitudes such as political trust when the political system does not offer real alternative policies to those who are dissatisfied.

Measurement and Operationalization

There have been different attempts to measure this concept. Several authors attempted to measure this concept with indicators of internal or external political efficacy, institutional trust, or even political cynicism. These attempts have mostly produced confusing and misleading results. Following the definition proposed above, measurement of political dissatisfaction should include indicators of approval and evaluations in terms of the distance between the individual's own policy or party preference and the policy alternative that the individual identifies with a particular party. Underlying here is the assumption that if an individual completely agrees with the policy alternative of a particular party as he or she perceives it (i.e., when a respondent places both himself or herself and the party at the same point on a scale), the individual is more satisfied with the party's policies than if the individual identified the party with a policy some distance from his or her own position on this issue. In this line of thought, Miller emphasized the role of policy evaluations and perceptions in trying to measure political dissatisfaction, creating an index based on policy preferences. However, as information on issue preferences is limited, costly to obtain, and not always very valid or reliable, scholars tend to use other general indicators related with governmental or incumbent approval and other political and economic evaluations. Some authors have even shown that the indicator of general satisfaction with the functioning of a democracy can be a good indicator of political dissatisfaction since it is highly contaminated by incumbent and government evaluations.

Mariano Torcal
Universitat Pompeu Fabra
Barcelona, Spain

See also Alienation; Legitimacy; Social Capital; Support, Political; Trust, Social

Further Readings

Citrin, J. (1974). Comment: The political relevance of trust in government. *American Political Science Review, 68*(3), 973–988.

Dalton, R. J. (2004). *Democratic challenges, democratic choices: The erosion of political support in advanced industrial democracies.* Oxford, UK: Oxford University Press.

Di Palma, G. (1970). *Apathy and participation: Mass politics in Western societies.* New York: Free Press.

Farah, B. G., Barnes, S. H., & Heunks, F. (1979). Political dissatisfaction. In S. H. Barnes, M. Kaase, et al. (Eds.), *Political action: Mass participation in five Western democracies* (pp. 409–447). Beverly Hills, CA: Sage.

Gamson, W. A. (1968). *Power and discontent.* Homewood, IL: Dorsey Press.

Gunther, R., & Montero, J. R. (2006). The multidimensionality of political support for new democracies: Conceptual redefinition and empirical refinement. In M. Torcal & J. R. Montero (Eds.), *Political disaffection in contemporary democracies: Social capital, institutions and politics* (pp. 46–78). London: Routledge.

Miller, A. H. (1974). Political issues and trust in government: 1964–1970. *American Political Science Review, 68*(3), 951–972.

Torcal, M., & Montero, J. R. (2006). Political disaffection in comparative perspective. In M. Torcal & J. R. Montero (Eds.), *Political disaffection in contemporary democracies: Social capital, institutions and politics* (pp. 3–20). London: Routledge.

Domestic Politics and International Relations

The study of the impact of domestic politics on international relations is currently experiencing a renaissance after a long period of neglect. Almost all work in international relations today deals in some way with the effect of processes within state borders on events outside of them. Scholars are interested, for instance, in how different interest groups and political parties formulate their foreign policy preferences, how institutions aggregate those different agendas, and how the combination affects strategic interaction among states and their ability to get what they want.

This entry provides a review of the recent contributions to the study of domestic politics in the international relations (IR) literature. As this literature is vast, the notion of paradigms of IR is used—that is, particular ways of looking at world politics that are premised on different key concepts and variables. Certainly, not all the research being done on the role of domestic politics fits neatly into one of these approaches, nor is it necessary for it to do so. However, paradigms do provide a way of categorizing the different ways in which domestic politics is understood by IR scholars to affect IR, and in doing so, they provide a fruitful tool for systematizing such research. First, however, a very brief history of the study of domestic politics in "modern" IR will be provided, covering roughly the past 30 years, noting the rise and fall and then the rise again of interest in domestic politics.

By a domestic politics approach, IR scholars generally mean analysis that tries to unpack the state into its constituent parts—that is, opening up the "black box" for closer examination. Consequently, scholars intent on demonstrating how states with different domestic institutions act differently in their IR can be said to use such an approach. This should be distinguished from those analysts who focus on the microlevel or individual level, such as the foreign policy role of key decision makers in the executive branch. This is the criterion used here to distinguish domestic politics from the individual- and systemic-level analyses.

A Brief History of the Study of Domestic Politics in Modern International Relations Theory

Scholarship aiming to understand how what occurs within a state affects its dealings with other states has a long intellectual pedigree, generally categorized under the rubric of "liberalism" in IR theory. Liberalism is sometimes understood as any argument or approach that disaggregates the state into its component parts. It finds its origins in the pluralism evident in liberal political theory in which individuals each pursue their own destiny. However, examining the processes occurring within a state is not confined to democracies. Autocracies have their own domestic political processes, although they are generally less obvious and transparent to the outside world. This approach to explaining world politics reached its apex with the publication of *Power and Interdependence* in 1977 by Robert Keohane and Joseph Nye. In it, they not only described an ideal type of world in which nation-states were disaggregated into their constituent parts—branches of government, bureaucracies, the media, advocacy, and interest groups—but they also argued that these parts coordinated with one another across country lines, sometimes in pursuit of a common agenda that put them at odds with other elements in their own societies. In this view, the state was not unimportant, but it was certainly viewed as becoming increasingly weaker in an increasingly interdependent world.

With what one might call the structural revolution in IR theory, approaches focusing on domestic politics were placed on the back burner for almost 20 years. Just after Keohane and Nye's book appeared, Kenneth Waltz's *Theory of International Politics* dealt an almost deadly blow to the more pluralistic approach to IR that includes a study of domestic politics. Waltz argued that only by understanding states as operating within a system could the persistence of patterns in world politics, most notably the pursuit of the national interest through the acquisition of military power, be explained. Due to the anarchic nature of IR, a characteristic of the international system, states of all different shapes and sizes pursue very similar policies. Given the lack of any enforceable law or governance in international affairs, states need to help themselves. Survival requires the accumulation of power and the subordination of other elements of state policy, such as trade or the pursuit of human rights, to that end.

Under this logic and given the stakes involved, domestic politics should not be allowed to intrude into the state decision-making process. Given individuals' shared interest in security within a nation-state, the system encourages them to come together in support of the national interest. This justifies what has been known as the "billiard ball" or "black box" approach, according to which, what happens inside the nation-state is seen as unimportant. For all intents and purposes, all states can be understood as being unitary actors responding to the same systemic pressures, with only their relative power left to distinguish them from other states. The implication is that it is the obligation of the state to guard the interests of society, even if this involves disregarding particular domestic interests. It should act as a unitary actor protecting a country's national interest, understood as more than just the aggregation of the preferences of individuals and groups within society. States should judge and interpret their environment as objectively as possible, since mistakes can have serious consequences.

This left little room for domestic politics. Only through systemic analysis can one understand the continuity in IR. By this, Waltz meant the big questions such as why phenomena such as war and the balance of power continually occur, which constitute the core of the discipline. Under Waltz's logic, domestic politics explains variation in foreign policy—that is, the smaller questions of day-to-day management of international affairs in contrast to international politics proper. Hence, domestic politics is a matter of foreign policy, not of IR. A generation of neorealist or structural realist scholars set out to show how the state might be understood as a unitary actor, coming together in times of threat to meet challenges in the rational way dictated by the system. The clear influence of this structural revolution in IR theory was reflected in the adoption of the unitary action assumption by

Keohane in his next critique of realist IR theory, *After Hegemony*. Keohane took issue with the pessimistic conclusions reached by structural realism but not with its foundational premises.

In recent years, however, there has been an explosion of interest in the influence of domestic politics in IR. This is partially the product of two new approaches, rationalism and constructivism, which question realist assumptions. However, even realists have begun to investigate domestic politics, although generally in keeping with the basic premises and constraints of structural realism. Certainly, not all research focusing on domestic politics explicitly identifies with these three approaches, or needs to. Given the enormous size of the literature, however, looking at the logic of these three strands helps us sort through and make sense of the broader trends in the discipline.

Rationalism

Although rationalism is a very broad approach that may not constitute a coherent paradigm of IR, it is nevertheless possible to identify central themes and a certain logic to the approach. It is based first and foremost on the presumption of utilitarianism on the part of individual political actors. Rationalism understands its objects of inquiry as calculating, egoistic players whose only possible bond is self-interest. Focusing on individual utility maximization leads rationalists naturally to break apart the state and look into the black box of structural realism.

While the premise of self-regarding rational action might strike one as somewhat trivial at first, it has been revolutionary for IR theory and has led scholars to question what was often the received wisdom of structural realism. For example, while a realist would stress how the importance of national security serves as a powerful incentive against petty domestic politics interfering with foreign and security policy, rationalists have focused on the often diverging interests within a country, for example, economic interests or office-seeking politicians. There is a large literature that argues that major questions of war and crisis initiation are dictated by the electoral calendar or the domestic popularity of the government, sometimes known as the "diversionary theory of war." Many have found an economic rationale underlying grand strategy—the fundamental and guiding principles of a state's foreign policy generally assumed to be driven by a state's structural position (i.e., its relative power or its geography) or its core ideology. In this view, instead of reflecting a national consensus or security imperatives, security policy is actually a post hoc rationalization of the economic interests of powerful constituencies seeking to secure and protect export markets.

By examining domestic politics in this manner, rationalists allow us to see that there is often a divergence between state policy and what is good for the society as a whole. This often makes for suboptimal policy. For instance, leaders might continue fighting wars long past the point at which victory is likely because they fear the domestic political consequences of defeat. The same counterintuitive logic is seen in international political economy, as states embrace protectionism that does not serve their interest in terms of collective welfare but rather serves a narrow parochial economic interest with a greater stake in the outcome.

Institutions are generally crucial for explaining this gap between the aggregate welfare of a nation-state and its actual policy. Depending on their configuration, institutions privilege some actors over others, giving a few the opportunity to direct policy in a direction that serves their parochial interests rather than the common good. This disjuncture, however, is something that varies. Democratic foreign policy is thought to reflect more accurately the aggregate welfare of society as a whole inasmuch as democracies have more inclusive representation. However, given the asymmetry of the stakes and information, there is often a wide gap between collective welfare and policy in democracies as well.

The sum total of this view is a very pluralistic approach of "boxes within boxes." Individuals aggregate into groups on the basis of common interests and engage in a process of bargaining and strategic interaction to realize their preferred ends. Domestic institutions are a key mediating influence as they privilege some groups over others. Once a particular policy preference emerges among competing groups, say political parties in a democracy, the broader group to which they both belong, say the state, then engages in a process of bargaining and strategic interaction with other like units.

Domestic institutions are also important in strategic interaction in that they allow for credible signaling in bargaining situations, a primary concern of rationalism given its utilitarian core. For instance, in the case of a deterrence dispute or a trade negotiation, a state must undertake an action that states that lacked resolve would or could not pursue. This can be done through "costly signaling." As the accountability of democratic institutions means that leaders face significant costs in case of failure, military escalation by democratic countries during crises is therefore a more credible signal of resolve than similar action by nondemocracies. Another option is to make private information about a country's commitment and capability more public to demonstrate that it has the power and the resolve to prevail. Democracies are also more transparent and information rich, allowing them to convey more persuasively their interests and will when resolute or their peaceful intentions when they want to reassure. However, the nature of democracies can also prove negative in negotiations. Given the presence of elections in democratic societies, negotiating partners might fear that democracies are less able to commit long term to their agreements, such as an agreement to liberalize trade.

Even while rationalists embrace the domestic level of analysis, its emphasis on institutions often comes at the neglect of domestic politics per se. Domestic institutions structure the political environment, acting as constraints on self-interested political actors. Rationalists focus on structural features, albeit at the domestic level, but do not devote great attention to the political process that plays out within them. As in the case of neorealism, the preferences of political actors are more often assumed than explained. A common assumption is that politicians are "office seekers" who design their political agenda, in foreign as well as domestic affairs, around the goal of obtaining or maintaining office. Using these universally applicable assumptions, rationalists can explain variation in terms of the differing types of domestic institutions. These are structural arguments that do not problematize the origin of political preferences and the political process that brings about policy. Interestingly enough, rationalism can be thought in some ways to apply structural realist assumptions at the domestic level of analysis. Politicians, like states, seek to survive, with survival meaning in this case staying in office and maintaining power. Rationalism does not truly eliminate the black box assumption. It merely adds boxes of different colors and sizes.

Constructivism

The constructivist research project is a powerful critique of structural realism. At its core is the idea that the nature of IR is social and as such is a constant project of construction. If foreign affairs are marked by mistrust, balance of power, and war, this is because states in the system have embraced a particular *culture of anarchy*, to use Alexander Wendt's term. Given the social nature of IR, however, it is capable of being reconstituted on the basis of different social practices. States can adopt different norms and identities that lead to different behaviors and transform the character of international interaction. The nature of anarchy is not determinate or given; it is mutable. Constructivism's logic leads to a natural focus on "ideational factors" such as norms, values, culture, and ideology.

While this opens up the potential for political agency and therefore domestic politics, the tendency among IR constructivists has been to focus on social interaction at the international level rather than the domestic level. The norms it uncovers and explains are generally systemic ones that lead to transnational convergence in behavior. The empirical fact that states of very different kinds adopt the same norms is a tool for showing, theoretically, that norms are powerful. Given that the primary task of the paradigm is to illustrate that what are regarded as immutable features of international politics, such as self-help or sovereignty, are, in fact, social constructions, this is not surprising.

Demonstrating that IR are social also involves showing that these cultures of anarchy become reified, accepted as a given on the part of those embracing the norms undergirding the system. That is, constructivists emphasize the structure in the mutual constitution of structure and agency. Constructivism's interest tends to be directed at these types of major macro outcomes or constitutive principles of the international system as a whole.

While this focus on reification and systemic norms is intrinsic to their efforts to show that norms matter, constructivists sometimes lose sight of why the social nature of IR means that there should be great domestic political contestation on issues of IR. Reification means the absence of agency, as political actors do not question but rather take these institutions for granted.

There are some exceptions, however. Constructivist literature has focused on the effect of a state's political culture, embodied in its political institutions. Democratic institutions lead to different types of foreign policy on the part of democracies not because of the incentives they provide to self-serving politicians but because they are the instantiation of a set of values externalized in foreign policy. Democracies build stronger alliances based on consultation, mutual identity, and trust.

Constructivists show that countries are marked by particular foreign policy cultures, an ideational consensus on the principles that guide a country's relations with others, and that these are generally based on a unique conception of self or identity. These cultures are not objectively given by a state's strategic setting but socially constructed on the basis of collective experience and the interpretation given to those events.

Both the focus on culturally revealing institutions and foreign policy cultures tends to imply political consensus at home. However, constructivists are also interested in ideological contestation over foreign policy, often expressed by political parties that serve as vehicles for different conceptions of a state's role in the world consistent with the set of values they represent in domestic affairs. Such considerations may affect a state's alignment choices, steering them in the direction of like-minded actors abroad. It facilitates analogies between domestic and foreign policy situations that allow political parties to develop unique stands on foreign policy issues that contrast with their domestic political competitors at home.

Neoclassical Realism

Ironically, some of the most innovative and interesting work on domestic politics and IR is being done by neoclassical realists, heirs to Waltz's legacy of structural realism. While Waltz stressed that the anarchic nature of the international system provided incentives to states to help themselves, he was careful to note that it was ultimately up to states themselves to understand their environment and heed its lessons. The system could only punish after the fact, not force states to act in a particular way. Waltz does not necessarily expect that all states will act in ways that suit the system's imperatives, only that those that do not will be punished. Domestic politics, by privileging parochial interests over that of the collective whole, can divert states from their optimal path. When they do so, the system will punish them. Waltz thus offers an observation and a caution—where we see excessive intrusion of domestic politics into foreign policy making, the state will not act as a unitary actor and foreign policy will be suboptimal.

Neoclassical realists demonstrate that foreign policy yields the best outcomes when the state acts as a coherent and unified political entity. For instance, strong states better able to extract resources from society are more powerful than those that cannot and better able to innovate militarily to meet security challenges by emulating the practices of the best performing states. More fragmented states, which resemble liberalism's pluralist model, are less adapted to this environment. When states are captured by parochial interest groups serving their own narrow agenda, states are led in disastrous directions.

Social and elite cleavages inhibit group feeling and cohesion, which makes states less able to act as unitary actors. These divisions reflect the prioritization of more parochial interests such as ethnic, cultural, ideological, religious, class, bureaucratic, regional, or party over the national interest. Governments presiding over fractured societies are weak and are unable to take the steps necessary to counter real threats. These factors interfere with the proper operation of self-help and the balance of power. Institutions also matter. Strong legislatures representing smaller sectional interests interfere with the making of a rational foreign policy by policy executives. Legislatures provide an access point into foreign policy for narrow constituencies that pull the settlement away from an ideal point that the state would otherwise prefer.

Unlike in rationalism, however, the state is a genuine entity with a life of its own that acts as the true representative of national interests. It serves as the perspicacious and sagacious agent of a country's true interests, threatened by smaller, sectional interests. Society might be powerful enough to sometimes trump the state, but the latter is generally not merely the reflection of aggregation of the former's interests, as is the case with the more pluralist theory of the state adopted by liberals. This concept of the state separates the neoclassical realist literature from others that stress how state representatives are primarily motivated by their own egoistic and parochial interests in staying in office. Where this does occur, neoclassical realism endeavors to show that there are severe consequences. The system punishes.

Current Trends

These three paradigms offer three unique ways in which processes at the domestic level of analysis influence IR. In rationalism, domestic politics are the battleground for competing individual and group interests over the state. In constructivism, the domestic level is the place where we locate culture and identity conceptions that guide a country's foreign relations in a unique way or a place in which this is contested on ideological grounds. In realism, domestic politics is the source of obstacles to a state's pursuit of the genuine national interest. Currently, rationalism is the starting point for most who take domestic politics seriously. Constructivism has dealt less with the subject. Ironically, neoclassical realists provide some of the most exciting new thinking on the topic.

Brian Rathbun
University of Southern California
Los Angeles, California, United States

See also Constructivism; Democratic Peace; Foreign Policy Analysis; International Relations, Theory; Liberalism in International Relations; Rational Choice; Realism in International Relations

Further Readings

Keohane, R. O., & Nye, J. S. (1977). *Power and interdependence*. New York: HarperCollins.

Milner, H. (1998). Rationalizing politics: The emerging synthesis of international, American, and comparative politics. *International Organization, 52,* 759–786.

Moravcsik, A. (1997). Taking preferences seriously: A liberal theory of international politics. *International Organization, 51,* 524–553.

Rathbun, B. (2004). *Partisan interventions: European party politics and peace enforcement in the Balkans*. Ithaca, NY: Cornell University Press.

Rathbun, B. (2008). A rose by any other name: Neoclassical realism as the logical and necessary extension of structural realism. *Security Studies, 17,* 294–321.

Risse-Kappen, T. (1995). *Cooperation among nations: The European influence on U.S. foreign policy*. Princeton, NJ: Princeton University Press.

Waltz, K. (1979). *Theory of international politics*. Reading, PA: Addison-Wesley.

Durkheim, Émile (1858–1917)

Émile Durkheim was a French sociologist who taught at the University of Bourdeaux and successively at the Paris Sorbonne. His main works include *De la Division du Travail Social* (The Division of Labor in Society, 1893), *Les Règles de la Méthode Sociologique* (Rules of Sociological Method, 1895), *Le Suicide: Etude de Sociologie* (Suicide: A Study in Sociology, 1897), and *Les Formes Elementaires de la Vie Religieuse* (The Elementary Forms of Religious Life, 1912). Durkheim committed his whole professional life, within the French university system, to the task of establishing sociology as a scientific discipline distinct from philosophy, psychology, and other social sciences. His intent (one might label it "positivistic") was to ground sociological discourse on an objective approach to *social facts* (i.e., collective "manners of thinking, acting, and feeling"). In pursuing this goal, he gave relatively little attention in his writings to expressly political phenomena, the intellectual discussion of which, in the fin de siècle French context, was chiefly the concern of philosophical, juridical, and ideological discourses.

It can plausibly be said that Durkheim's work was chiefly aimed at providing a scientific solution to the problem of social order. Aware as he was that the problem had been posed in the most compelling manner by Thomas Hobbes, a political philosopher, Durkheim did not accept Hobbes's own, expressly political solution of that problem: the fear-motivated submission of all individuals to an unchallengeable sovereign brought into being artificially by the social contract. According to Durkheim, humans are *by nature* intensely social beings, and the individuals' ability and willingness to coordinate their reciprocal activities could not be reliably and durably sustained by the coercive enforcement of the sovereign's commands. Normally, instead, it expressed a condition of solidarity between individuals.

However, in the course of social evolution, solidarity, while always called on to ground and guarantee the social order, presents itself in varying forms. In *Division* (the first and perhaps the greatest of his works), Durkheim sharply simplified such variation by subsuming all forms of solidarity under one or the other of two basic, sharply contrasted types: mechanical and organic.

Mechanical solidarity is associated with the earliest stages of social evolution. Here, societies that might be called "primitive" have small populations operating over relatively large territories (thus, characterized by low densities). Their subunits differ very little from one another, enjoy rather poor command over natural resources, and are sustained only by rudimentary technology and by minimal division of labor; thus, only to a very limited extent do they exchange their products with one another. In turn, within each of these subunits, the individuals themselves, the ultimate components of the society, again barely differ from one another, for the existence of all is managed with reference to universally shared bodies of practical knowledge. They all harbor within their minds the same cultural patrimony, which all subunits periodically revisit and celebrate on ceremonial occasions. Thus, in a society of this nature, solidarity rests on the extent to which individuals are like one another, for they all share understandings of what is true and proper, which cover all the contingencies of the society's simple, tradition-based mode of existence.

The second type of solidarity, the organic, is common to all other societies, which may be labeled "advanced." These have evolved out of "primitive" ones and are characterized chiefly by a more or less advanced degree of division of labor, and thus by their complexity. Typically, these societies have large populations, densely settled over the territory, and their subunits are differentiated from one another; part of the population, for instance, inhabits towns; the rest inhabits the countryside. Different regions specialize in the production of different goods and services and exchange these with one another. Furthermore, the individuals inhabiting each locale are in turn differentiated from one another with regard to their occupation, their skills, and the techniques they employ in performing their specialized tasks. Inevitably, they share only a small portion of the increasingly wide, diverse, and changing cultural patrimony of society as a whole while they make their own and identify with a narrow subset of it that varies according to their differentiated conditions. Put otherwise, individuals are increasingly individualized and expected and authorized to operate as self-conscious, self-activated, and self-seeking units.

The solidarity typical of such a society can no longer be grounded on the universal sharing of beliefs, values, and norms; rather, it rests on the extent to which society's subunits (and within these the individuals themselves) have become differentiated, for this process makes them interdependent and generates wide and flexible networks of exchange. Individuals, in particular, are induced to relate to one another largely by means of contractual arrangements, where each party's legitimate concern to foster its own well-being finds expression in its efforts to provide for the other's needs and to contribute to the other's well-being.

One could say that according to Durkheim, in the course of social evolution, the basis of solidarity (thus, the solution to the problem of social order) shifted from *cultural homogeneity* to *material interdependence*. Neither basis, on the face of it, is expressly political in nature. However, Durkheim's profound, sophisticated sociological discourse could not fail to confront at different points, in a more or less explicit and elaborate manner, the political dimension of social affairs, and as it did so, it produced results worthy of attention, including some of marked originality.

Perhaps Durkheim's awareness of political phenomena was inspired by some traits of his personality. He was an austere man and rather forceful man, whose strong sense of hierarchy and authority found expression, among other things, in the way he selected and trained his intellectual followers and monitored their performance. Also, in the pursuit of his lifelong mission—again, establishing the discipline of sociology in the French academic system—he found it useful to associate himself in the capacity of a privileged consultant with the ministerial authorities who managed that system with a maximum of central direction and control from above. He took the same approach in attempting to accomplish a broader task—having sociology (which to him meant *his own* sociology) provide an intellectual inspiration for the structuring of the educational system at large and for the shaping of its programs specifically. Finally, he was a committed republican, convinced of the superior merits (not just for France) of the "principles of 1789"—*égalité*, *liberté*, *fraternité*—the last of which inspired his concern for solidarity.

In any case, Durkheim also acquired a sense for the significance of political phenomena from the substance of his own sociological discourse, beginning with *Division* and other early works. According to *Règles de la Méthode Sociologique*, for instance, only "manners of thinking, acting and feeling" that are "sanctioned" can be qualified as proper "social facts." That is, there have to be arrangements, no matter how varied across time and space, for rewarding those observing some manners (positive sanctions) and/or for taking to task those disrespecting and violating them (negative sanctions). Now, the idea of sanction itself, and especially of negative sanction, imparts, so to speak, a vertical dimension to the collective unit (a society or a group), which privileges some "manners" and makes them its own; it suggests, however, a structured and justified command-and-obedience relationship between one part of the unit and the others, and such a relationship has traditionally been seen as intrinsically political.

In *Division*, Durkheim used as an indicator of the changing nature of solidarity the laws and in particular the prevailing nature of the sanctions applied to violators of legal expectations. In primitive societies (where the laws were mostly customary), *punitive* sanctions prevailed, for society's intense reprobation of universally shared and cherished norms found expression in the infliction of suffering on violators. In advanced societies, punishment became less significant than *restitutive* sanctions. Here, whenever a party to a contract fails to fulfill its obligations toward the other, it allows the latter, after fulfilling its own obligations, to recover, in some way or other, what it legitimately expected. Again, this way of conceptualizing the critical difference points up, however implicitly, the vertical structuring of both societies; for even the application of restitutive sanction, no less than that of punitive ones, requires some authoritative, power-*ful* arrangements for identifying and dealing with violators. And it seems appropriate to think of such arrangements as political in nature.

Furthermore, in *Division*, Durkheim opposes the excessive emphasis placed by some authors (both in his time and ours) on contract (and thus on the market) as *the* distinctive, superior way of coordinating the actions of individuals in advanced society. Durkheim agrees that contract is central to such a society and is associated with significant moral values (such as individual autonomy and responsibility) but insists that it should not displace authoritative, public institutions as regulators of private, self-interested activities. For him, "not everything is contractual about the contract." Individuals can enter binding contracts only by availing themselves of the *institution* of contract, a set of preexistent authoritative arrangements establishing who can make contracts about what matters, in what forms, and with what consequences.

The notion of institution is of utmost significance for Durkheim, who at one point defined sociology itself as the science of institutions. These are arrangements (bodies of personnel, distinctive practices, resources, etc.) focused on matters of central concern to the whole society and backed by a diffuse moral appreciation of them as morally valid. But one can also see institutions as operating "vertically"—operating on society, as it were, from above.

This aspect of institutions is apparent when Durkheim raises political matters—developing and applying expressly sociological insights and eschewing any involvement in the politics of parties—and proposes institutional changes he obviously hoped would be realized via legislative acts.

The characteristic dynamism of modern society, propelled by the competition between economic actors and by technological innovation, often generates conflicts and crises and atomizes individuals, depriving them of the normative guidance previously provided by tradition and by local memberships. The machinery for regulation operated by the state and activated by the competition of parties mostly cannot deal effectively with these new circumstances. To remedy this, the state itself should empower "corporations"—public bodies identified with specific sectors of the economy—to monitor closely the economic process, discipline the unruly relations between employers and employees, and instill in both a sense of responsibility toward customers and toward the public at large.

In a lecture course Durkheim taught repeatedly at Bordeaux and at the Sorbonne, whose content was made available after World War II, this *corporatist* proposal (put forward also in other writings) complements a sophisticated and original argument on expressly political affairs. According to an insight already present in *Division*, a critical step in the advance of the division of labor was constituted by the emergence of a *part* of society entrusted with the task of controlling the *whole* by making and implementing specific decisions concerning its relations with the (natural and social) environment.

In modern society, this task is performed in a particularly purposeful and self-conscious manner by the state. This stands to the rest of society as the brain stands to the organism of advanced animal species; its essential function is thus the acquisition and processing of information about the relevant environments, leading to the formation and implementation of collective decisions.

This bold, "cybernetic" view of the state connects in Durkheim's views with a particular understanding of democracy. Durkheim considered democratic arrangements for political representation via parties, coupled with bureaucratic administration, as particularly effective ways of generating and accumulating information on social affairs and communicating it upward and downward, thus allowing decisions to be made in a particularly deliberate and sophisticated manner. Thus, according to Durkheim, democracy was a particularly apt mechanism for, as it were, putting and keeping state and society in touch with one another. By the same token, democratization was the most appropriate political expression of the advance of modernity.

Although here Durkheim confronts expressly political themes, and imaginatively emphasizes their informational and communicational aspects, he articulates, once more, a "vertical" understanding of them, attending chiefly to political institutions and paying little attention to what could be considered the "horizontal" aspect of political experience—that is, to *politics*. This may express, again, his prevalent concern with social order, which induced him to consider with distaste social, and thus also political, conflict.

Gianfranco Poggi
University of Trento
Trento, Italy

See also Corporativism; Democracy, Theories of; Functionalism; State

Further Readings

Filloux, J.-C. (1977). *Durkheim et le socialisme* [Durkeim and socialism]. Geneva: Droz.

Lukes, S. (1937). *Emile Durkheim: His life and work*. London: Allen Lane.

Mueller, H.-P. (1983). *Wertkrise und gesellschaftsreform: Emile Durkheims schriften zur politik* [Value crisis and social reform: Emile Durkheim's writings on politics]. Stuttgart, Germany: Enke.

Mueller, H.-P. (1993). Durkheim's political sociology. In S. Turner (Ed.), *Emile Durkheim: Sociologist and moralist* (pp. 95–110). London: Routledge.

Poggi, G. (2000). *Durkheim*. Oxford, UK: Oxford University Press.

E

ECONOMIC POLICY

Economic policy is one of the central activities of government. This entry describes the changing nature of economic policy, discusses the important forms of government intervention, and describes the emerging trends in public sector interventions in the economy. The decision makers, content, and policy instruments of economic policy have changed substantially over the past 50 years. It was once a more straightforward matter to define economic policy than it is in the early 21st century. At one time, economic policy was about macroeconomic policy—the targeting of policy objectives such as growth, employment, and inflation through measures to influence the demand side of the economy principally using the policy instruments of fiscal policy, taxation, and government spending. National governments were the principal decision makers and were held accountable by their electorates in democratic societies for the conduct of economic policy. In communist societies, the economic plan was a central device for achieving socialism.

Internationalization of Economic Policy

At the domestic level, governments now share their economic decision-making authority with central banks and a range of bodies that seek to regulate the financial sector in particular. However, economic policy has become much more internationalized with a complex system of multilevel decision making. Bodies such as the World Trade Organization (WTO) set the rules within which economic activity takes place, while the G-8 and, more recently, the G-20 seek to coordinate the economic policies of major governments. The 27 member states of the European Union (EU), particularly those using the euro, face another regional level of decision making that shapes that economic policy.

This more internationalized system of decision making reflects the increasing internationalization of economic activity. In the period between World War I and World War II, nation-states pursued policies of autarchy in which they aimed for self-sufficiency either within their own national borders or in economic systems that embraced their colonies. These policies of protectionism, combined in some countries with the denomination of their currency in gold—the gold standard—which produced overvalued currencies, undermined international trade and the efficiency gains it could provide through comparative advantage.

After World War II, the world, under the leadership of the United States, sought to move toward a more liberal international economic system, at least within the Western bloc. The so-called Bretton Woods institutions and agreements—the International Monetary Fund (IMF) and the World Bank, along with the General Agreement on Tariffs and Trade (GATT)—sought to create the economic and political space within which this system could operate. In the period between the end of World War II and the 1970s, this system was built around fixed exchange rates, pegged against the dollar backed by gold and by relative capital immobility. These arrangements were not without costs for the United States, but it was often prepared to bear them in the pursuit of broader policy objectives such as underpinning Western economies in the context of the Cold War. Through this period, international trade was stimulated by the reduction of tariffs through the GATT negotiating rounds, although one unintended consequence was that nontariff barriers to trade started to assume a greater importance.

This system came under increasing strain in the late 1960s. In part, this was a consequence of the long-run decline in the economic dominance of the United States, exacerbated in the short run by the economic strains arising from the Vietnam War and the costs of programs, including Medicare, enacted as part of President Lyndon Johnson's Great Society program. In the early 1970s, the Bretton Woods system collapsed when the United States removed the link between the dollar and gold. There was no longer a fixed rate system but only currencies that fluctuated substantially against each other. One consequence was that the European Economic Community started moving toward a zone of regional currency stability, initially through the creation of the exchange rate mechanism of the European Monetary System in 1979. Another was the need for a system of economic policy coordination between the major economic powers to compensate for the inability of the United States to take a leadership role by the establishment of economic summits in 1975 that became institutionalized as the G-7 (later G-8) process.

The industrial dominance of the United States and the leading European countries was being challenged by the emergence of newly industrializing countries (NICs). The first of these countries was Japan. The economic success of the United States had been built on Fordist systems of the mass production and consumption of standardized economic goods, which relied on securing efficiencies that drove down unit costs to produce price-competitive goods that could first be sold to the large American market and then through branch plants overseas. A major motive for establishing the common market in Europe was to achieve comparable economies of scale there.

Japan had initially competed with the West by producing relatively standard industrial goods such as steel and ships using Western technology. Among the first motor vehicles that Japan sold to the United States were pick-up trucks. Over time, however, Japan came to increasingly emphasize quality in production and produce more sophisticated goods that challenged Western goods in their home markets. As Japan moved upmarket, a number of countries, initially "Asian tigers" such as South Korea and Taiwan, moved into more standardized goods production using "off-the-shelf" technology. They were able to sell these goods into Western markets because of the lowering of tariff barriers. This led to the development of new policy instruments in the form of nontariff barriers to restrain the penetration of Western domestic markets by these NICs.

The sharp increase in oil and other commodity prices in the early 1970s triggered a process known as stagflation. Since 1945, the Western countries had enjoyed a long postwar boom with high growth rates and low levels of unemployment, although this was accompanied by increasing problems in restraining inflation. This had been made possible by a variety of factors: technological breakthroughs (e.g., the widespread use of new materials such as plastics), postwar reconstruction, demographic shifts, and, more controversially, new forms of economic policy. In the 1970s, growth slowed, unemployment rose, and inflation increased. These problems were compounded by the second oil crisis or shock at the end of the 1970s, which triggered a recession in the early 1980s.

Information Technology and Economic Policy

Toward the end of the 1980s, the world began to experience the benefits of an industrial revolution comparable to the harnessing of steam power and the development of the factory system at the end of the 18th century and the science-led industrial revolution at the end of the 19th century based on the application of electricity to industrial production and the development of science-based industries such as chemicals. This third industrial revolution was based around information technology. Mainframe computers had been available for some time and had been used for tasks such as processing inventory and wages. However, it became possible to network micro or desk computers, and this opened up new possibilities in information management in companies. Advances in computer-aided manufacturing and computer-aided design also facilitated automation and precision production in manufacturing. This in turn made possible more variations in forms of production, thus facilitating differentiated batches of production to serve niche markets. The Fordist world of production was to become a post-Fordist one, in terms of both systems of production and consumer preferences.

By creating new, faster, and more reliable systems of communication, information and communication technology (ICT) also facilitated the increasing integration of the world economy. This had been a long-term process, driven by technology, market organization, and policy. It increasingly became referred to as "globalization," although this was a highly contested term. It implied that there had been a step change in the organization of the world economy, and not all analysts accepted that this had been the case. There was also considerable normative dispute about whether globalization was beneficial or not, and this conditioned policy responses to it.

Trends in Economic Policy

It is possible to identify a general trend in the world economy with the following characteristics. Except in intermittent recessions, trade tended to grow faster than national output, and this upward trend line was invariably resumed once a recession had ended. Foreign investment tended to grow faster than trade as multinational companies became more dominant in the international economy. Indeed, trade increasingly took the form of trade within multinational companies as they integrated production. The development of these multinational companies created a need for a more integrated international financial system with a wider and more sophisticated range of financial instruments. These financial instruments increasingly developed a momentum of their own with a particular development of trading in debt. Regulators seeking to supervise these developments faced asymmetric information problems: There were too many instruments to supervise effectively with the resources available, particularly given their complexity, and in any case the prevailing philosophy was one of "light touch" regulation. The risks associated with this approach became apparent in the financial crisis of 2008, which was in large part triggered by the issuance of subprime mortgages to borrowers with poor credit histories. These subprime mortgages were then bundled into packages of debt that were traded.

The fundamental problem was that economic policy was still conducted largely at a national level, yet the economic structure was increasingly internationalized. What institutions were available to coordinate economic policy at an international level, and how effective were they? The IMF had been created after World War II to resolve the problem of global imbalances. These problems remained in the first decade of the 21st century, although now the imbalance was principally between the G-2 countries of China and the United States, an economic relationship that was complicated by tensions over security and human rights issues. In economic terms, China played a major role in funding U.S. government debt, enabling the United States to run both a budget deficit and a balance of payments deficit, the latter enabling China to sell its exports into the U.S. market.

The IMF had a clearer role in the era of fixed currencies and relative capital immobility in the period up to the early 1970s. After that, its role became less certain, although it was able to make a significant intervention when British economic policy ran into trouble in 1976, imposing restraint on public expenditure. However, as the story of those events has become better known as the archives opened up, it has become apparent that domestic actors were using IMF intervention to impose greater restraint than was strictly necessary, in part as an effort to restore government authority, which had been undermined by the role of the trade unions as veto groups on economic policy. As larger sums of money started to move around the international financial system, the sums available to the IMF became less significant as leverage on government policy. In the financial crisis of 2008, the IMF was able to bail out smaller economies in serious trouble, such as Iceland and Latvia, but was not able to come up with any general solution to the world's economic problems.

That role would have at one time fallen on the G-7, which became the G-8 with the addition of Russia. In the 1970s, the G-7 had a clear focus on economic issues, but it increasingly became a general summit covering topics such as terrorism. Even when it did make pronouncements on economic issues, such as international trade negotiations, it seemed to have little impact. The leaders, particularly the host country, seemed to treat the event as a major profile-raising exercise for domestic audiences, and if any effective work was done, it was by the "sherpas"—officials who prepared the ground for the meetings through bilateral and multilateral contacts. Moreover, the G-7/G-8 appeared to be increasingly irrelevant to the structure of the world economy. It did not include China and other emerging economies. The G-7 included Canada and four EU member states: Britain, France, Italy, and Germany. The inability of the EU to have one representative demonstrates its lack of unity on economic issues, with Britain, in any case, outside the eurozone. The EU's initial response to the financial crisis was sluggish, with no clear central policy making on economic policy, despite the increasing prominence of the Economic and Financial Council (Ecofin). Perhaps the most important set of policy instruments, certainly in the context of a recession—namely, those of fiscal policy—remained under the control of the individual member states.

The international response was to establish a new body, the G-20, which included emerging countries such as Brazil, China, and India, reflecting the shift in balance of economic power in the world. Some commentators were concerned that the new body had gone too far in the direction of inclusiveness at the expense of effectiveness by bringing in countries such as Mexico, Saudi Arabia, and South Africa. Nevertheless, it made a relatively effective start to its work, eclipsing the G-7/G-8. This body was left with an uncertain future role. However, it was important not to exaggerate the extent to which power had shifted away from the traditional power holders. Even if the United States could no longer impose its preferred policies on others, an economic policy solution that did not have U.S. support was unlikely to meet with much success.

The international architecture of economic policy making has become increasingly important, but the domestic level remains significant. It is important to appreciate how domestic policy has evolved over time and policy instruments have changed. This in turn needs an appreciation of the importance of economic ideas, the way in which economic orthodoxies have changed, and the significance of economists, within government, in business, in the media, and as academics in influencing policy.

Although the precise origins of the term *macroeconomic policy* are the subject of some dispute, it was not in common use before World War II. This was because governments did not seek to influence the macroeconomic aggregates beyond pursuing policies of sound money through a balanced budget and a stable currency. In the 1930s, governments in the United States, Britain, and elsewhere did become involved in seeking to rationalize and revive depressed economic sectors, but these were sector-specific policies.

World War II led to a step change in the involvement of governments in economic management, particularly as a consequence of a widespread commitment to the maintenance of full employment, although what exactly constituted full employment was left ambiguous. In his *General Theory* (1936), John Maynard Keynes had set out an approach that sought to run the economy at a higher equilibrium employment level. Because Keynes wrote so extensively, and his work was interpreted and developed by his disciples after his death, there is some controversy about what exactly may be called "Keynesianism." In essence, it advocated the use of fiscal policy measures to avert high levels of unemployment by stimulating investment and consumption by reducing taxes and/or increasing public expenditure. Spending on public works was seen to be especially effective, because through the "multiplier" effect originally identified by Richard Kahn, an initial investment could have a far greater effect on levels of economic activity. To prevent the economy from "overheating" as a result of productive capacity not keeping pace with growing aggregate demand (as, e.g., in an economic boom), taxes could be increased and public expenditure reduced. The existence of an inverse relationship between unemployment and inflation rates, as described by the Phillips curve, offered a range of choices to policymakers. Keynesian ideas spread more rapidly in some countries than in others, but by 1971, even an economic conservative like President Nixon was able to state to an ABC News reporter, "I am now a Keynesian in economics."

These new techniques of economic management seemed to be associated with a prolonged period of relatively high growth and high levels of employment after World War II. However, it is difficult to disentangle how much of this good performance was due to Keynesianism and how much due to other factors such as the deployment of technological advances made during the war, the needs of postwar reconstruction, and demography. Keynesianism also brought problems in its train. It tended to produce stop–go cycles as the economy was accelerated and decelerated. This was particularly a problem in Britain, and it was thought that one solution was the adoption of the French indicative planning model, but a watered-down version was applied in the United Kingdom.

As Keynes himself had admitted, he had no solution to the inflation problem in a full-employment economy. Countries resorted to various forms of income policies, either agreements between centralized employers' organizations and trade unions, which held back wage growth, or interventions by government to impose restraint. The more voluntary agreements, provided they were negotiated by centrally controlled trade unions, tended to be more effective, especially in countries such as Austria and Sweden, and came to be known as "corporatism" in the academic literature, although many policymakers did not like this term because of its fascist associations. These arrangements worked well enough in a period of rapid growth when no one left the bargaining table empty-handed, but they became more challenged when growth rates fell in the 1970s, which led to the phenomenon known as stagflation: low growth and high inflation (and increasing unemployment). Barry Eichengreen argues that the institutional arrangements that worked so successfully up to the 1970s were less successful afterward and slowed down the ability of the corporatist economies to readjust to changed conditions.

The apparent failure of Keynesianism led to a search for a new economic orthodoxy, and this was found in the work of Milton Friedman and his advocacy of monetarism. Even in Britain, money supply indicators proved to be too volatile and unreliable to provide a decision rule for the economy, but they ushered in a period when there was a substantial reduction in government involvement in the economy. Controlling inflation was prioritized over maintaining employment, and the accepted economic orthodoxy was reversed so that microeconomic, supply-side policy was seen as the mechanism for achieving growth, while macroeconomic policy focused on the control of inflation. Monetary policy instruments, principally the control of the interest rate, became more important relative to fiscal policy instruments, and emphasis was placed on targeting the inflation rate as the principal economic objective.

One consequence was an increase in the importance of independent or quasi-independent central banks. The Bundesbank as a constitutionally guaranteed autonomous actor had been central to the success of postwar German economic policy. In Britain, from 1997 onward, the New Labour Government had handed over operational control of interest rates to the Bank of England. The reality of "all power to the central bankers" was reflected in the importance of the Federal Reserve to U.S. economic policy. In the EU, the European Central Bank faced a disjuncture in the eurozone: It controlled monetary policy, but fiscal policy remained the responsibility of the member states. The Stability and Growth Pact was supposed to restrain excessive deficit spending by member states but in practice failed to do so as it was defied by France and Germany.

Colin Crouch and others have traced the evolution of a system of "privatized Keynesianism" in which an economic stimulus was provided, particularly in Britain and the United States, by large amounts of consumer and corporate debt. This increasingly became linked to the housing market, with Britain using an inflation measure as a policy guide that omitted the cost of housing. However, the provision of increasingly risky loans and the increased selling of financial instruments made up of packages of debt undermined the foundations of the financial system. It led to the subprime crisis in the United States triggered by defaults on loans to poor-quality borrowers, which developed into a serious global financial crisis and contributed to the onset of a severe economic recession by 2008.

A range of new policy instruments had been deployed to cope with this crisis, with renewed efforts to achieve effective international coordination. Some saw the events as vindicating Keynesianism, although not all countries, particularly Germany, were enthusiastic about a fiscal stimulus based on government debt. New techniques of "quantitative easing" involved central banks buying up corporate and government debt, so the central banks remained key actors alongside government economic ministries. The result of these measures was a massive increase in public debt. Although low interest rates reduced the cost of servicing it, taxes would have to be raised and public expenditure cut in order to repay it.

These events triggered a renewed debate about the role of financial services regulation. Although there was a general view that financial markets regulation needed to be tighter, it was less clear how that might be achieved. Some argued that risk was central to a market economy and reducing it too much would stifle the economy, whereas others argued that reckless risk taking had been encouraged and needed to be curtailed. After a period when economic policy had been "depoliticized" and automatic rules supervised by central bankers had replaced political discretion, economic policy was suddenly no longer a matter of purely technical debates but moved to the center of the political stage once again.

Wyn Grant
University of Warwick
Coventry, United Kingdom

See also Budgeting, Rational Models; Central Banks; Political Economy; Public Budgeting

Further Readings

Eichengreen, B. (2007). *The European economy since 1945*. Princeton, NJ: Princeton University Press.

Fourcade, M. (2009). *Economists and societies: Discipline and profession in the United States, Britain and France, 1890s to 1990*. Princeton, NJ: Princeton University Press.

Gamble, A. (2009). *The spectre at the feast: Capitalist crisis and the politics of recession*. Basingstoke, UK: Palgrave Macmillan.

Grant, W. (2002). *Economic policy in Britain*. Basingstoke, UK: Palgrave Macmillan.

Hall, P. (Ed.). (1989). *The political power of economic ideas*. Princeton, NJ: Princeton University Press.

Walker, A., & Sen, G. (2009). *Analyzing the global political economy*. Princeton, NJ: Princeton University Press.

ECONOMIC STATECRAFT

Economic statecraft is the use of economic means to pursue foreign policy goals. Foreign aid, trade, and policies governing the international flow of capital can be used as foreign policy tools and are considered the most common forms of economic statecraft. In principle, policies governing the international movement of labor could also be considered instances of economic statecraft if they are intended to promote foreign policy goals; but such measures are not usually included under the rubric of economic statecraft and are not considered here. This entry discusses the forms of economic statecraft, the uses of such instruments, and approaches to the study of economic statecraft.

Economic techniques of statecraft are distinguished from other foreign policy tools, such as the following: (a) military statecraft, which concerns the use or threat of military force; (b) diplomacy, which concerns negotiation; and (c) propaganda, which concerns manipulating verbal or visual symbols. Most foreign policies consist of some combination of these techniques.

Economic statecraft takes many different forms, including both positive and negative sanctions. Negative sanctions are actual or threatened punishments, while positive sanctions are actual or promised rewards. Examples of negative sanctions include the following: refusing to export (embargos), refusing to import (boycotts), covert refusals to trade (blacklists), purchases intended to keep goods out of the hands of target countries (preclusive buying), deprivation of ownership (expropriation), punitive taxation, aid suspensions, and asset freezes. Examples of positive sanctions include preferential tariffs, subsidies, foreign aid, investment guarantees, and preferential taxation of foreign investment.

Neither the study nor the practice of economic statecraft is of recent origin. Although the Athenian use of the Megarian Decree may be the most famous example from ancient times, it was surely not the first. Examples of various types of economic statecraft can be found throughout recorded history. The use of economic means to pursue foreign policy goals has been discussed by a number of thinkers through the ages, including Plato, Aristotle, John Locke, Francis Bacon, Montesquieu, David Hume, Adam Smith, Immanuel Kant, Alexander Hamilton, Friedrich List, John Stuart Mill, Woodrow Wilson, and John Maynard Keynes.

Tools of economic statecraft have been used to pursue a wide variety of foreign policy goals.

These include preparing for war, preventing war, fighting a war, promoting democracy, punishing human rights violators, promoting communism, opposing communism, promoting economic development, discouraging economic development, preventing regime change, encouraging regime change, and many other goals. Pericles, Theodore and Franklin Roosevelt, Woodrow Wilson, Adolf Hitler, Joseph Stalin, Winston Churchill, and the United Nations (UN) have used various forms of economic statecraft to pursue goals that were sometimes noble and sometimes nefarious. Like other techniques of statecraft, economic tools can be used wisely or unwisely, justly or unjustly, depending on the situation.

In the first part of the 20th century, the League of Nations generated hopes that warfare could, to some extent, be prevented or replaced by economic sanctions. The League of Nations imposed sanctions on Italy in response to its invasion of Ethiopia in 1935 partly to punish Italy but also to warn Hitler of its members' determination to resist aggression. The failure of the sanctions with respect to both goals generated a widespread belief that such measures do not work. Much of the last half of the 20th century was dominated by the acceptance of this belief as conventional wisdom.

Answering the question of whether economic sanctions "work," however, is more complicated than it seems. There is not even a consensus on what "work" means in this context. Does it mean complete achievement of the primary goal? Of all goals? At what cost? And in comparison with which alternative techniques? These are just some of the questions begged by the misleading question: "Do economic sanctions work?"

The first—and most important—step in evaluating the utility of any technique of statecraft, including economic sanctions, is identifying what goals were being pursued with respect to which targets. In the case of the League of Nations sanctions against Italy, for example, impressing Hitler was probably more important than stopping Italian aggression. In addition, there was a desire not to impose such hardship on Italy that it might undermine the fascist regime and bring the communists to power.

Human beings in general—and nation-states in particular—rarely, if ever, pursue only one goal at a time with respect to only one other individual or group. When a country imposes economic sanctions on another country, it is usually pursuing multiple goals of varying degrees of importance with respect to a multiplicity of other actors in the international arena. Although no assessment of the overall success of an influence attempt based on economic means could be expected to consider all goals and targets, it is reasonable to expect such assessments to consider the most important goals and targets. Most attempts to assess the success of economic sanctions, however, consider only one goal with respect to one target.

An additional complication is that success is almost always a matter of degree. In statecraft, as in everyday life, complete success in goal attainment occurs rarely—if ever. The potential number of degrees of success (or failure), of course, is infinite; thus, any attempt to measure degrees of success must involve simplification. Whether a 3-point scale or a 16-point scale is more appropriate is a matter for reasonable dispute, but a dichotomous conception, which fails to allow for any degree of success, is difficult to justify. Most influence attempts are likely to result in some degree of success with respect to at least some goals and some targets.

In addition to estimating goal attainment, determining the success of economic sanctions involves estimating the costs of the undertaking. Measuring the costs of economic sanctions is usually easier than measuring the noneconomic costs. Political costs are likely to be both important and difficult to measure. This does not mean that reasonable estimates are impossible.

Perhaps the most difficult step in assessing the utility of economic techniques of statecraft is setting such assessments in the context of the logic of choice. Without comparable evaluations of the costs and benefits of alternative techniques of statecraft, estimates of the likely costs and benefits of economic sanctions are of little or no interest. For policymakers, it is the relative utility of a policy option that matters. This is especially important when the alternative to economic sanctions is military force. In such cases, economic sanctions often provide a policy alternative with modest expectations of goal attainment at relatively low cost, while military force involves a higher probability of goal attainment accompanied by much higher costs.

If the goal is to choose the policy alternative with the most utility, the rational choice may be one that accomplishes fewer goals at much lower cost. The U.S.–led invasion of Iraq in 2003 provides an example. The economic sanctions against Iraq were not providing the United States with the degree of assurance that Iraq was not building weapons of mass destruction that it desired, and sanctions were not bringing about regime change even though several of the goals specified by the UN had been accomplished. Advocates of war argued that the costs of maintaining the sanctions could go on for years, whereas the costs of war would be limited in both time and magnitude. As is often the case, the costs of war were significantly underestimated. Thus, to make a policy-relevant assessment of the utility of economic sanctions, it is necessary not only to estimate the costs and benefits of sanctions but also the costs and benefits of alternative courses of action, such as war.

The critics of economic sanctions often dismissed them as "merely symbolic," implying that they were empty gestures aimed at giving domestic audiences the false impression that "something was being done." Although some economic sanctions may well fit this description, not all do. As scholarly work in foreign policy and international politics has come to recognize the importance of signaling, the symbolic uses of economic sanctions have received serious scholarly attention. Economic sanctions can serve as costly signals that increase the credibility of foreign policy stances. Thus, the symbolic use of economic sanctions can be an important way to influence the actions and attitudes of other countries. Game-theoretical models have been especially useful in the study of this aspect of economic statecraft.

In sum, economic statecraft is the use of economic means in pursuit of foreign policy goals. As with other foreign policy tools, economic statecraft can be used to make either threats or promises and to either punish or reward. And as with other foreign policy tools, success is often difficult to evaluate.

David A. Baldwin
Princeton University
Princeton, New Jersey, United States

See also Diplomacy; Foreign Aid and Development; Foreign Policy Analysis; International Political Economy; International Trade; Sanctions

Further Readings

Baldwin, D. A. (1985). *Economic statecraft*. Princeton, NJ: Princeton University Press.

Baldwin, D. A. (1999/2000). The sanctions debate and the logic of choice. *International Security, 24,* 80–107.

Hufbauer, G. C., Schott, J. J., Elliott, K. A., & Oegg, B. (2008). *Economic sanctions reconsidered* (3rd ed.). Washington, DC: Peterson Institute.

Knorr, K. (1975). *The power of nations: The political economy of international relations*. New York: Basic Books.

O'Sullivan, M. L. (2003). *Shrewd sanctions: Statecraft and state sponsors of terrorism*. Washington, DC: Brookings Institution Press.

Economic Theories of Politics

Economic theories of politics use theoretical approaches from economics to analyze political phenomena. This entry examines assumptions about rational choice that are shared by the two fields and discusses the application of rational choice approaches, including game theory, to the analysis of individual choice and strategic interactions. It then considers the way in which social choice theory can illuminate the aggregation of individual choices in collective decision making. The relationship of such processes to voting, particularly in the context of democracy, is described. Finally, the entry looks at the application of economic theories to international relations with respect to issues such as arms races, war, development, and trade.

Rational Choice and Strategic Interactions

Most political scientists would agree that politics involves "control, influence, power, or authority." If one adds Max Weber's concerns about government, legitimacy, and the state, together with Aristotle's more normative focus on issues of participation, citizenship, and justice, one has a fairly complete picture of what Robert Dahl calls the "political aspect." One can see immediately how politics touches every dimension of human activity, including the procedural or distributional dimension (who gets what, when, how, why, and at whose cost?); the legal or statist dimension, which involves issues of governance and legitimacy; and the ethical or normative dimension, which revolves around questions of citizenship, justice, and participation.

The study of politics, like economics, also involves preferences, interests, and trade-offs. But unlike economics, where the emphasis is on scarcity and efficiency, in politics the primary emphasis is on power, influence, and authority, with strong ethical and normative overtones, concerning justice, membership, and citizenship. In a free market, the allocation of scarce goods and resources takes place according to the logic of the marketplace (the price mechanism)—that is, the interaction of supply and demand. The exercise of power, however, takes place in the ideational, legal, and institutional confines of political systems.

Then, what have economic theories added to the study of politics? Political scientists know that politics, unlike economics, is not interested narrowly in the allocation of scarce goods and resources. Although politics affects markets through policies, laws, and rules that regulate competition, in a mixed capitalist system politics is not directly concerned with the individual economic decisions of consumers and producers or the optimal allocation of scarce resources. But politics, like economics, does involve choices and strategic interactions. This is where those who advocate a positive approach to the study of politics join forces with economists to lay the microfoundations of political analysis. So-called rational choice approaches in political science share common assumptions with economics about human rationality and strategic decision making, and economists seek to construct economic theories to explain political behavior.

The most common economic theories of politics take rational choice approaches. They assume that individuals are rational in the sense that they will make choices to "maximize their chances of achieving their goals" (Barbara Geddes, 2003, p. 177). They give priority to agency (individual rational actors) over structure (institutions and other political constraints). They assume that individuals have goals and that institutions and other factors affect individual strategies and preferences. In this framework, utility-maximizing individuals will do what they can to achieve their goals, engaging in strategies to anticipate the actions of others (their opponents), who will in turn anticipate the actions of the other side. Strategic interactions therefore refer to the ways in which each individual not only looks out for his or her own interests but also takes into account the interests and strategies of others. In this rational choice framework, conflict, cooperation, and the give and take of political life are outcomes of myriad strategic interactions.

Rational choice approaches often use game theory to understand the complexity of strategic interactions in situations of conflict. Developed by applied mathematicians in the mid-20th century, game theory is widely used by economists and, to an increasing extent, by political scientists. Game theory is not a "theory" in the sense of a set of claims, laws, or propositions about the way the world works. It is rather a method of constructing theories, and it offers the analyst a set of concepts and tools that enable him or her to formalize arguments. Game-theoretic analysis requires careful specification of the beliefs, wants, and needs of individuals and a clear understanding of what strategies are available to them. The need for specificity makes game theory less useful as a tool for applied political and social science; nonetheless, it helps political and social scientists understand the logic and structure of politics, whether they are studying voting, international relations, or policy making.

Strategic Interactions and Collective Choice

Rational choice approaches or "economic theories" of politics begin with assumptions about individual rationality. But how do collective decisions relate to individual choices? To answer this question, social choice theory focuses on how individual preferences are aggregated and what roles institutions play in "engineering" social (or collective) choices. The disjuncture between individual and collective preferences is summed up in Condorcet's paradox—the puzzle that calls into question the notion of majority rule, which is incapable of producing a stable relationship between individual preferences and collective decisions. Kenneth Arrow sought to address the puzzle of how individual preferences affect collective choices and concluded that there is no mechanism short of a dictatorship that can achieve collective rationality. This is known as Arrow's impossibility theorem.

As an economic theory of politics, Arrow's theorem helped explain how democracies work. Arrow does not prove that collective rationality—which would guarantee the aggregation of individual preferences transitively—requires concentration of power in the hands of a single decision maker or dictator. The main implication of Arrow's theorem is that institutions are *the* critical link in understanding how radically divergent individual preferences are translated into collective action in rather stable ways.

Building on Arrow's social choice theory, William Riker (1980) argues that

> politics is *the* dismal science because we have learned from it that there are no fundamental equilibria to predict. In the absence of such equilibria we cannot know much about the future at all, whether it is likely to be palatable or unpalatable, and in that sense our future is subject to the tricks and accidents of the way in which questions are posed and alternatives are offered and eliminated. (p. 443)

While Riker recognized the failure to find equilibria for collective rationality, he stressed the importance of institutions for the smooth functioning of a democracy. His answer to the question of how democracies can reach collective decisions despite the lack of equilibria was

> [in a democracy] the way . . . tastes and values are brought forward for consideration, eliminated, and finally selected is controlled by . . . institutions. And institutions may have systematic biases in them so that they regularly produce one kind of outcome rather than another. (p. 443)

Following Riker's expansion of Arrow's theorem, political and social scientists must ask how institutions matter in democratic decision making. Many studies have debated whether and how institutional structures determine the existence and location of equilibria for collective choice. Anthony Downs points out that governments are not really interested in maximizing individual voters' preferences but in maximizing votes. In his analysis, the sole point of politics is to gain and hold power; in a two-party system, politicians are forced to take positions as close as possible to the median voter. Thomas Romer and Howard Rosenthal (1978), however, showed that once agenda control is taken into consideration, the collective outcome of policy making is often not the median voter's ideal point. Kenneth Shepsle (1979), in a more generalized way, shows that institutions could induce equilibrium for collective choice where individual preferences alone would not yield equilibrium.

Voting has long been a principal subject of political analysis, from the so-called paradox of voting, first identified by Downs—why should individuals bother to vote when the costs of voting (time and effort) far outweigh the benefits (the likelihood that a single vote will influence the outcome)?—to strategic voting. Maurice Duverger points out that the strategic behavior of voters and candidates is heavily influenced by electoral institutions. Duverger's law holds that plurality voting in single-member districts tends to produce two-party systems, whereas voting based on proportional representation or multimember districts leads to multiparty systems and coalition governments. In two-party systems, candidates have an incentive to obfuscate and avoid taking strong stands on key issues as they jockey for position vis-à-vis the median voter; in multiparty systems, however, candidates have an incentive to take stronger positions to attract a significant minority of voters, but their positions may change once they enter a coalition government. In retrospect, these arguments may seem obvious or almost self-evident, but they are in fact early examples of economic reasoning in the study of politics.

Following Duverger and others, one can see that political parties and electoral systems are the most important institutions for the smooth functioning of a democracy. They translate and aggregate individual preferences into policy. John Aldrich (1995) argues that democracies would not work without parties. He shows how parties regulate the number of people seeking office and how they mobilize voters to achieve and maintain the majorities needed to implement policy once they have gained power. By aggregating individual preferences, parties help solve (or at least side step) Arrow's impossibility theorem, and they move society toward equilibria that make governing possible. This argument for the importance of parties relies heavily on economic reasoning, and it assumes that voters are able to make informed choices.

To take a concrete example, the founders of the American Republic believed that voters would have enough information to make the "right choices." In the *Federalist Papers*, James Madison argued that people should have the "virtue and intelligence to select men of virtue and wisdom." Studies of democratic elections, however, have found that individuals lack detailed information about politics. Voters make their choices by trusting expert opinion, by accepting consensus judgments, and by using what little information they have to make their choices at the ballot box. In short, voters vote *as if* they are well informed and using the limited information available to them.

Strategic Interactions and International Relations

Economic theories of politics abound in the study of international relations. At first glance, it is not surprising to see that states are more often than not in conflict, because each state must pursue its own interests, maximizing its power and wealth to provide security. States are trapped in a security dilemma; if they do not correctly assess threats from other states, they risk their very survival. The logic of the security dilemma is often explained in game-theoretic terms, using the so-called prisoner's dilemma, whereby states are compared with two prisoners, accused of a crime, jailed, and isolated by prosecutors. Each prisoner is offered a deal if he or she will confess. If both confess, each will receive a moderate sentence; if neither confesses, both will receive a reduced sentence; and if one confesses and the other remains silent, the one who confesses goes free while the other receives a life sentence. The smart play is to keep quiet, but by not knowing what his or her partner in crime will say to the authorities, each prisoner's temptation to confess (defect) is overwhelming. The point is that actors have an individual incentive to defect, which leads to an outcome of mutual defection even though both would be better off with mutual cooperation. With respect to the security situation, it may seem irrational for states to engage in an arms race, thereby increasing the propensity to go to war and making it more difficult to resolve their differences through negotiations. However, military expansion yields a better outcome than disarming when the other side arms, even though the best outcome would be for both states to disarm. Thus, the states engage in an arms race, which is costly and may lead to war.

Motivated by this puzzle, James Fearon (1995) offers several "rationalist explanations for war." He postulated three mechanisms by which conflict can lead to war. First, war may occur if there is uncertainty about the adversary's capabilities, such as the size of the military, the effectiveness of military technology, the quality of leadership, and the contribution, if any, of allies to the cause. There may also be uncertainty about the adversary's resolve to fight wars, which raises questions about how much each side values the "good" that is in dispute and what the ultimate cost of war will be in terms of blood (casualties) and treasure (wealth). Second, Fearon argues that war can occur if states are unable or unwilling to honor bargains—versions of the prisoner's dilemma. If there are advantages in striking first, if the relative power of one side is expected to grow rapidly, or if the bargaining is over goods that affect future bargaining power, then states may face a choice between war today on favorable terms and the threat of war tomorrow on unfavorable terms (the shadow of the future). Finally, bargaining may not help avoid war if the good in dispute is indivisible.

Hence, Fearon's economic theory of war shows us that a mutual (and rational) preference for peace is not sufficient for states to overcome the "asymmetric information problem." As a result, states generally have incentives to exaggerate and misrepresent their capabilities and their resolve to fight—recall the bluster of Saddam Hussein in the run-up to the 2003 invasion of Iraq by the United States (also known as the second Gulf War). In such a situation, the danger is that brinksmanship—the act of sending credible signals about capabilities or resolve—may raise the risk of war.

One of the enduring issues in the study of international relations is the so-called democratic peace argument, with many competing arguments about why democracies rarely, if ever, fight wars against one another. One explanation is that democracy increases the political costs of war by making elected leaders accountable to people who ultimately must pay the costs of war. Another explanation is that democracy increases transparency, which may help overcome the asymmetric information and commitment problems in strategic interactions between states, thereby reducing the severity of the security dilemma. An alternative argument is that democratic states never go to war with each other because they almost always have a common enemy—an example would be the Soviet Union versus the Western democracies during the Cold War period.

Some have argued that international institutions reduce the risks of war and help alleviate the security dilemma. The United Nations (UN), for example, provides a set of rules that structures international relations in such a way as to promote collective security. However, the UN, like other international institutions, lacks the capacity to enforce its rulings. States have the right to defend themselves when attacked—the self-help rule—but there is no guarantee that other states will come to their aid. Still, some would argue that international institutions mitigate the effects of anarchy and promote cooperation. By setting up norms and rules and through creating repeated interactions (iterative games), international institutions help states overcome the security dilemma, and cooperation rather than conflict becomes the norm of international relations—this is the so-called shadow of the future. Institutions offer a set of norms and standards of acceptable behavior; they provide information about activities of other states (potential adversaries), thereby creating linkages across issues and raising the costs of free riding and defection (Lisa Martin & Beth Simmons, 2001).

Economic theories of politics not only help us understand the classic issues of war and peace in international relations, but they are essential to understanding trade policy. If economists are correct in arguing that trade is beneficial for all trading partners, then why do some states insist on protecting their markets, restricting trade, and damaging the prospects for world economic growth and welfare? Is this not irrational behavior? The theory of comparative advantage shows that trade is beneficial for society as a whole, but it is known that trade creates winners and losers. Using the simple economic logic of the Stolper-Samuelson theorem, which postulates that protectionism leads to an increase in the income of a state's scarce factor, Ronald Rogowski (1989) developed a theory to explain how political coalitions form to support or oppose free trade. The logic of collective action suggests that when the benefits of protectionism are concentrated and the costs are diffuse, the state may be captured by powerful interests who stand to win from protectionism. Conversely, if a state pursues free trade and opens its markets, the benefits will accrue to the exporting sectors, in which case the state is able to overcome collective action problems and pursue a trade liberalization policy. This is a classic example of an economic theory of politics, which helps us understand the politics of trade.

The same reasoning can be applied to problems of development. At a meeting of the World Trade Organization (WTO) in Cancun in 2003, talks aimed at reducing protectionism and liberalizing trade broke down due to a conflict between rich and poor states. Here, one can see how the decline of U.S. economic power has made it more difficult to achieve multilateral trade agreements. In the absence of a leader (hegemon), there is a greater propensity for states to pursue beggar-thy-neighbor policies and to free ride—gain the benefits of trade but without paying the costs (lower tariffs and less protectionism) needed to maintain an open international economy. More important, the integration of developing countries into the global market has generated a backlash against trade liberalization. As Geoffrey Garrett (1998) notes, unskilled workers in developed countries—losers from free trade according to the Stolper-Samuelson theorem—have opposed globalization because of fear of downward pressure on wages, while people in developing countries (and the activists who speak for them) fear vulnerabilities to global markets. Though economists argue that social welfare policies to support the losers of free trade and the recent trend of globalization (such as income compensations) are more efficient than trade protectionism, the formal model of Daren Acemoglu and James Robinson (2006, chap. 10) shows that global capital mobility makes it more difficult for a democracy to provide social welfare policies that the majority of the citizens may want.

Conclusion

Supporters of economic theories of politics, specifically rational choice approaches, argue that it is the most scientific (hypothetico-deductive) way of studying politics, and as such they contend that it is high time that the discipline of political science embraces a more rationalist and utilitarian theory of politics. Some would go so far as to say that rational choice theory constitutes a "scientific revolution" à la Thomas Kuhn, moving the study of politics away from its formal-legal and social-psychological roots and in the direction of more systematic and falsifiable propositions. The irony is that economics seems to be moving in the opposite direction, with a renewed emphasis on socio-psychological approaches to the study of markets and economic behavior. Perhaps the two disciplines will meet halfway and a true political economy will emerge; but this seems unlikely because the objects and the subjects of inquiry are quite different. This entry cannot begin to resolve the dispute between rationalists and social psychologists and institutionalists; instead, it falls back on Weber who leaves ample room for rationalist and interpretivist approaches to the study of politics.

James F. Hollifield and Hiroki Takeuchi
SMU
Dallas, Texas, United States

See also Collective Security; Game Theory; Rational Choice; Social Choice Theory; Strategic (Security) Studies

Further Readings

Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. New York: Cambridge University Press.

Aldrich, J. H. (1995). *Why parties? The origin and transformation of political parties in America*. Chicago: University of Chicago Press.

Fearon, J. D. (1995). Rationalist explanations for war. *International Organization, 49,* 379–414.

Garrett, G. (1998). *Partisan politics in the global economy*. New York: Cambridge University Press.

Geddes, B. (2003). *Paradigms and sand castles: Theory building and research design in comparative politics*. Ann Arbor: University of Michigan Press.

Martin, L. L., & Simmons, B. A. (2001). *International institutions: An international organization reader*. Cambridge: MIT Press.

Riker, W. H. (1980). Implications from the disequilibrium of majority rule for the study of institutions. *American Political Science Review, 74,* 432–446.

Rogowski, R. (1989). *Commerce and coalitions: How trade affects domestic political alignments*. Princeton, NJ: Princeton University Press.

Romer, T., & Rosenthal, H. (1978). Political resource allocation, controlled agendas, and the status quo. *Public Choice, 33,* 27–43.

Shepsle, K. A. (1979). Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science, 23,* 27–59.

Effectiveness, Bureaucratic

This entry discusses bureaucratic effectiveness and its meanings. These various meanings include an agency's political capabilities and effectiveness at acquiring resources, its ability to preserve its autonomy, and ultimately its ability to produce public value and to measure that value. The validity of metrics, however, is related to clarity of goals, and some goals and functions are inherently clearer than others. In addition, political controversy makes clear goal definition problematic.

Varieties of Bureaucratic Effectiveness

It is difficult to find a single definition of bureaucratic effectiveness. A big part of the explanation for that lies in the many distinct functions that bureaucratic agencies have. Some have very defined customer service characteristics. An automobile licensing agency, for example, may be able to calculate the wait time for its customers and seek to cut that through better management. Taxing agencies might well calculate their effectiveness in terms of recovering revenue. But what, some have asked, does a foreign ministry do and how does one measure whether or not it is meeting its goals, assuming it could define them?

It is necessary to distinguish between bureaucracies in the private and public spheres. Generally, private firms measure their success, if they are public stock corporations, in the confidence of their investors. This is typically reflected in the value of its stock price and its trajectory. In the public sphere, however, bureaucratic agencies' missions are defined by legal and political authority, and usually these missions are articulated in ways that are highly general, aspirational, vague, and potentially even contradictory.

Bureaucracies may be effective or ineffective across many different dimensions. Some dimensions consider their prowess in politics and gaining power, others with maintaining their autonomy and organizational culture, and yet others with procuring resources sufficient for the tasks to which they have been assigned. Additional dimensions have to do with defining and gaining acceptance for favorable accounts of success, gaining reputation for innovation and efficient management, and other measurements of organizational outputs. And sometimes bureaucratic agencies have the good (or bad) fortune to have outcomes move in a favorable (or unfavorable) direction despite their own activity. The perceived quality of police agencies, for example, is often related to the rise and fall of crime rates whether or not the police actually are directly responsible for these trends. Ironically, the higher the crime rate, the more the police may be valued whereas the lower the crime rate, the more the public and its politicians may find downsizing the police acceptable. This produces perverse incentives for the police agencies, although they are not the only agencies for whom less demand threatens their resources.

Resources and Alliances

As individuals, we generally prefer more to less. Organizations are no different. They too prefer having greater resources to work with than fewer. For one thing, it makes it more likely that agencies can achieve more of their goals than not. It is also a sign of an agency's power, raising the stakes for those who may be trying to cut it down to size. Having more resources rather than fewer also helps maintain morale and makes it more possible to retain highly qualified employees.

There is a need to distinguish, however, between resources that are mandated under law, such as pension benefits, and discretionary or appropriated resources. Transfer payments, for example, may account for a high level of mandated expenditures, whereas increasing agency personnel or giving more discretion as to how personnel may be used is a discretionary, and highly valued, resource.

Ensuring the availability of slack resources (sufficiency plus a reserve) is often an indicator of an organization's political prowess. Resource availability or lack thereof is also a function of the priorities of political leadership outside the organization. In view of these shifting priorities from the outside, the ability to generate support and allies of a longstanding nature and to be buffeted from political turbulence is the mark of a politically effective organization. Effective engagement with stakeholders and sources of political support and the ability to make its functions seem indispensable are essential to the political effectiveness of bureaucracy.

Capability and Instrumentation

No organization wants to be caught lacking the capability or instruments to do effectively the jobs for which it is tasked. Bureaucracies with capital and labor-intensive functions—notably in defense, transportation, infrastructure, and other high-capital investment sectors, including science and technology—often have two competing incentives. One is technical efficiency, defined here as the most effective route for the best marginal return on resources. Another is political support, emphasizing the extent to which capital projects can spread largesse sufficient to attract and sustain political support or be able to provide a compelling narrative. Sending astronauts to Mars, for example, provides a more compelling narrative than sending robots, though the scientific merits of doing so are at least debatable. Nearly all public agencies and especially those that can disperse capital projects face the choice between efficiency and political support. It is necessary to distinguish between the instruments and capabilities that are needed to effectively perform an agency's tasks and those needed to attract substantial political support. Too much emphasis on the first likely will lead to insularity and even irrelevance. But excessive emphasis on the latter frequently leads to the displacement of goals and certainly to inefficiencies.

Preserving Organizational Definition

All bureaucratic organizations face conflicts between their missions and the political needs related to their budgets and resources. This conflict can be viewed as a struggle, on the one hand, between the desire for organizational autonomy and securing the integrity of an organization's mission from outside interference and, on the other, the need to gain broad legitimacy in order to sustain organizational functions. As organizations accept more functions and diversify their stakeholders, they run the risk of losing control over mission definition. Not all bureaucracies desire to absorb new functions for precisely these reasons, but others, perhaps most, prefer more to less and will adapt their missions to gain political support.

Typically, bureaucracies dealing with finance, treasury, and overhead functions in financial accountability have been able to minimize external clientele interference in their affairs and even, to some degree, political interference. Some specialized scientific and technological functions as well as similarly situated military or intelligence units have also managed to preserve significant autonomy partly because their activities seem sufficiently esoteric and technical. The ability of bureaucracies to sustain autonomy depends on creating an aura of professionalism and success on matters that everyone cares about. An organization whose initial leadership has played an outsized role in establishing an aura of invincibility and indispensability can help sustain substantial organizational autonomy—and resources. From an agency's standpoint, the combination of resources and autonomy is the goal.

Innovation and Adaptation

So far, the discussion of bureaucratic effectiveness has focused on political problems—procuring support, gaining resources, ensuring that capabilities exist for the functions that have been assigned, and preserving organizational autonomy. These are, however, means to ends. From an agency's standpoint, those means provide advantages in power struggles and enable the bureaucracy to control its destiny. Public bureaucracies, however, are supposed to do more than harbor resources. Presumably, their purpose is to achieve public goals efficiently.

Thus, another way of examining organizational improvements is to assess how well a bureaucracy innovates or adapts to new conditions. Nonincremental change does not come naturally to bureaucracies weighted down by inertia. Both experts and leaders may account for organizations' adaptive and innovative capabilities. Technical experts within bureaucracies can play an important role in policy innovations. Organizational leaders may emphasize the importance of continuous adaptation. Slogans abound to call attention to leaders who claim to find ways to innovate, for example, "doing more with less." Whether there is actual innovation or mere cost-cutting is another question. Organizational leaders acquire a reputation by appearing to be innovative, and that becomes useful currency to the leaders, if not to their organizations. The language and symbols of innovation, however, usually outpace its actual occurrence.

Innovation is itself a challenging concept in the literature on bureaucratic effectiveness. It is not easily defined. One thing that can be said about an innovation is that it represents a parametric shift in the way in which some activity is conducted or in the way it is defined. The standard assumption is that it is better to innovate than enervate.

From Goals to Metrics

Increasingly, bureaucracies have to operationally define their goals and assess their progress through quantitative indicators. These procedures may be related to an organization's actual goals, but the goals may be redefined so as to be measurable. To the extent that metrics are outcomes related rather than outputs related, there likely are many uncontrollables present.

Policy areas especially subject to extensive analysis of metrics are education, health, and crime. An analysis of outcomes requires distinctions to be drawn between manipulable matters such as program design, or what agencies should emphasize, and nonmanipulable matters, such as demographics and social backgrounds. Policies may be responsible for only a small part of positive or negative change. Bureaucratic innovations to improve operations and outputs can be taken only within the existing policy parameters. Ultimately, assessing bureaucratic activities through metrics requires agreement on goals and on what the metrics should be. In areas such as education, debate continues on what the purpose of education is and how narrowly or broadly it should be pursued.

There is no doubt that there is more self-assessment undertaken by bureaucratic organizations and more assessments demanded by leadership. Performance evaluation is ongoing, and as with any measurable activity, much depends on the validity of what is being measured and how the outcomes can best be assessed. All assessments are imperfect, but that is still better than operating blindly. As agencies demand impact assessments in some cases from those whom they regulate, their own regulations also have become increasingly subject to cost–benefit analytics.

The most difficult element to measure in an agency often is management. Outside experts and agencies have collaborated in the United States to try to assess organizational goals for each federal agency. It is probably not terribly surprising to learn that less controversial agencies tend to do better at clarifying their goals or that agencies with higher proportions of civil servants among their senior executives tend to be managed better. Agencies that are politically controversial have more difficulties defining goals since these are likely to change with the political context. These agencies also tend to have a higher proportion of politically appointed officials. In most governments, some agencies are regarded as too important to fail, and party politics is minimized. But equally in most systems, other agencies are seen as offshoots of the policies of current governments or administrations, and some are seen as indulging in pork barrel politics.

In sum, performance evaluation is an imperfect, but necessary, science, and there is simply no question that organizations are being required to know more about potential problems, continually assess improvement efforts, and evaluate the costs of alternatives. This is the essence of a scientific-rational model of effectiveness evaluation in bureaucracies. But metrics demand clarity, and it is easier to distill an appropriate set of metrics when there is clarity about goals. That tends to push against the world of politics, which in turn limits bureaucratic effectiveness.

Bert A. Rockman
Purdue University
West Lafayette, Indiana, United States

See also Bureaucracy; Civil Service; Performance; Performance Management

Further Readings

Gormley, W. T., & Weimer, D. (1999). *Organizational report cards*. Cambridge, MA: Harvard University Press.

Hatry, H. P. (2007). *Performance measurement: Getting results* (2nd ed.). Washington, DC: Urban Institute Press.

Lewis, D. E. (2008). *The politics of presidential appointments: Political control and bureaucratic performance*. Princeton, NJ: Princeton University Press.

Selznick, P. (1949). *TVA and the grass roots: A study in the sociology of formal organization*. Berkeley: University of California Press.

Wolfe, P. J. (1993). A case survey of bureaucratic effectiveness in U.S. cabinet agencies: Preliminary results. *Journal of Public Administration Research and Theory, 3,* 161–181.

Efficacy, Political

Political efficacy is one of the most important topics in research on political attitudes and political culture. A large number of terms such as ego strength, self-esteem, self-reliance, subjective or civic political competence, or (lack of) powerlessness/futility can be found in the relevant literature as equivalents. The concept, which was first introduced into empirical research by Angus Campbell, Gerald Gurin, and Warren Edward Miller (1954), referred to the citizen's self-perception as a knowledgeable, active, and self-confident participant in political life. According to Campbell et al., the sense of political efficacy refers to the feeling that individual political action has or can have an impact on the political process. Political efficacy includes cognitive and behavioral components. The cognitive dimension refers to an individual's subjective belief that he or she is able to understand what is going on in politics, while the behavioral component taps the citizen's conviction that he or she can play an active role in politics, exert political influence, and thus cause political leaders and the political system to eventually react responsively to the citizen's actions. The following describes the development and measurement of the concept and then turns to the way in which political efficacy is used in political research.

Development and Measurement

In the first stage of research, political efficacy was regarded as a unidimensional concept important to the explanation of electoral participation. A cumulative scale to measure feelings of political efficacy consisting of five items was developed and included in the American National Election Survey (ANES) of 1952. The items were worded as follows:

1. "I don't think public officials care much what people like me think."
2. "The way people vote is the main thing that decides how things are run in this country."
3. "Voting is the only way that people like me can have any say about how government runs things."
4. "People like me don't have any say about what the government does."
5. "Sometimes politics and government seem so complicated that a person like me can't really understand what's going on."

Rejecting Items 1, 3, 4, and 5 and accepting Item 2 indicated feelings of political efficacy.

The second stage of research was characterized by refining the theoretical concept of political efficacy gradually and improving the measurement. From a theoretical perspective, Robert E. Lane's study *Political Life: Why People Get Involved in Politics* (1959) was an important step forward in the analysis of political efficacy. Lane regarded the prevailing unidimensional concept as too simplistic because it neglected to distinguish the individual's political self-image as a political actor from the individual's orientation toward the political environment. Accordingly, he proposed a distinction between the two separate dimensions of internal and external efficacy. The first component, internal political efficacy, entailed the belief that the normal citizen is able to understand and influence politics. External political efficacy referred to the perceived openness and responsiveness of the political system and the political elites.

Lane himself did not validate the distinction between internal and external efficacy. But his contribution served as a starting point for a broad and intense methodological debate on the meaning and the measurement of the concept. Several studies published from the mid-1970s on dealt with the validity of the concept and the quality of the items developed to measure feelings of political efficacy. Research carried out in this context followed two different lines: on the one hand, validating the distinction of internal from external efficacy and, on the other hand, improving the quality of the items used as measures of these two facets of the concept.

A first important contribution to the validation of the concepts of internal and external efficacy was made in an article published by Balch in 1974. In a construct validation based on external criteria, Balch showed convincingly that internal and external efficacy correlated with different indicators of political attitudes and behaviors: While internal efficacy items were related to political interest, knowledge, and conventional forms of political participation, external efficacy was closely linked to trust in government and participation in unconventional political activities, particularly civil disobedience. The construct of internal efficacy was represented by the "Voting is the only way" and "Politics is too complex" items, while "Officials don't care" and "People have no influence" turned out to be measures of external efficacy. One of the original items ("Vote decides how things are run") was dropped from the ANES surveys after 1952.

During the 1970s and 1980s, efficacy items were included in a considerable number of surveys. The improvement of the data situation encouraged a considerable number of methodological studies aimed to validate the concept further. They not only developed and tested additional items but they also used more sophisticated strategies of validation, particularly structural equation models. All studies confirmed the theoretical and empirical relevance of the distinction between internal and external efficacy. However, they did not convey clear results regarding the quality of the items and their relationship toward the constructs of external and internal efficacy. To improve the quality of the scales, two additional measures of external efficacy were included in 1968 in the ANES: (1) "Generally speaking, those we elected to Congress in Washington lose touch with the people pretty quickly" and (2) "Parties are only interested in people's votes but not in their opinions." These latter two, together with the item "Public officials don't care," were mostly used to measure external political efficacy, while the remaining three items served as indicators of internal efficacy. These six items also became part of opinion surveys conducted outside the United

States and in several cross-national surveys. In most recent studies, the "voting" item is no longer used as a measure of internal efficacy due to its ambiguity but replaced by some new items better suited for measuring the citizens' cognitive capability and feelings of competence. The distinction of these two latter subdimensions of feelings of internal efficacy was also confirmed empirically by several German studies published in the 1990s.

Use in Political Research

The concern with developing appropriate measures played an important, if not dominant, role in research on political efficacy. However, the concept was also applied in several fields of substantive research on peoples' political attitudes and behaviors. In the initial stages of research, feelings of political efficacy were not so much regarded as a research topic in its own right but rather as an explanatory variable in analyses of political participation. It was assumed and confirmed by empirical research that feelings of political efficacy were positively related to electoral turnout and other forms of political participation, even if education and political interest were held constant. In some of these national and cross-national empirical studies, feelings of (internal) political efficacy were introduced as independent explanatory concepts, and in others, they were considered as being part of a broader class of attitudes, labeled as "psychological involvements in politics." The traditional assumptions on the direction of the causal link of (internal) political efficacy to political participation were challenged by some empirical studies in the 1980s and 1990s, which showed a reciprocal relationship between political efficacy and participation. Accordingly, feelings of political efficacy not only encourage people to become active participants, but active participation in political life also consolidates and strengthens the citizens' self-perception as competent and influential political actors.

Most of the studies mentioned so far referred to conventional forms of political participation. As political protest became more widespread in the Western world, some attempts to explain the new phenomenon also referred to political efficacy. Some studies emphasized the interaction of political distrust and political efficacy as determinants of different types of political behavior. Accordingly, the coincidence of low efficacy and low trust was assumed to be a source of extreme political disengagement, because people characterized by a likewise combination of attitudes did not feel sufficiently competent and at the same time doubted that government officials would behave in an appropriate way if in the face of civic activity. On the other hand, a combination of high efficacy and low trust was seen as leading to political protest, civil disobedience, and sometimes even political aggression or violence, because people felt influential but did not trust the government to behave appropriately. People who primarily relied on conventional or reformist styles of political influence were characterized by a strong sense of political efficacy going along with strong trust in politics. Explanations of this type became familiar under the heading of the "efficacy–distrust hypothesis." However, empirical support for this assumption was rather weak.

Apart from research on political participation, the analysis of civic self-consciousness developed as an independent field of research in the 1950s. Scholars such as Gabriel Almond and Sidney Verba, Robert Lane, Paul Sniderman, and Alex Inkeles introduced political efficacy and related concepts as parts of the democratic personality or of a broader syndrome called "civic culture" or "participant citizenship," which was considered an important cultural prerequisite of a modern, particularly democratic, political community. In their study *The Civic Culture: Political Attitudes and Democracy in Five Nations*, Almond and Verba (1963) introduced subjective civic competence as an indispensable element of a democratic political culture. According to them, the civic culture is characterized by a pattern of political attitudes contributing to two different requirements of democratic politics: the power to govern effectively on the one hand and responsiveness to the people's demands on the other. While they conceived political trust as the attitude providing power to elites to make and implement authoritative decisions and policies, subjective political competence functions as a counterbalance to power in their view. Only competent citizens become active to make their voices heard in political life. The mere perception of the potentiality of civic activity—and less the activity in itself—was regarded by Almond and Verba as a sufficient condition of elite and systemic responsiveness. The idea of a necessary balance of political power and responsiveness (structure), and political trust and efficacy (culture) is emphasized by several theoretical and empirical contributions to research on political culture and democratic citizenship. Other contexts in which political efficacy or concepts similar to political efficacy can be found are studies of political apathy and political alienation. In research on political apathy (or lack of political involvement), lack of political efficacy is analyzed together with low political interest and political knowledge, and sometimes political participation. Research on alienation distinguishes between several components of the respective syndrome. Feelings of powerlessness, which are the most broadly analyzed component of alienation, are measured by efficacy items in most instances.

Oscar W. Gabriel
Universität Stuttgart
Stuttgart, Germany

See also Alienation; Apathy; Civic Culture; Civic Participation; Electoral Behavior; Political Culture

Further Readings

Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations.* Princeton, NJ: Princeton University Press.

Balch, G. A. (1974). Multiple indicators in survey research: The concept "sense of political efficacy." *Political Methodology, 1,* 1–43.

Campbell, A., Gurin, G., & Miller, W. E. (1954). *The voter decides.* New York: Wiley.

Prewitt, K. (1968). Political efficacy. In D. L. Sills (Ed.), *International encyclopedia of social sciences* (Vol. 12). New York: Macmillan.

Vetter, A. (1997). *Political efficacy: Reliabilität und Validität* [Political efficacy: Reliability and validity]. Wiesbaden, Germany: Deutscher Universitätsverlag.

Election by Lot

The election of public officers by lots, or lottery, is a procedure that can prevent the formation of a permanent leading group and diffuse knowledge of public affairs among the members of the community. It can be an appropriate formula in settings in which an assembly of members or representative council makes decisions by broad consensus or unanimity, and public jobs do not require high technical skills. The selection of delegates by turns and the subsequent rotation of people in public offices can have about the same effects as lotteries. There is an old tradition of choosing public officers by drawing lots that can be found in ancient and medieval local democracies, modern private settings, and some international organizations. In most cases, it goes together with the central role of the assembly to make decisions on the most relevant issues, typically by consensual agreement, on the assumption that the identification of a common interest should not be too difficult a task.

The most relevant historical experience of selection of delegates, representatives, or public officers by lots was developed in Athens during the democratic period from the mid-5th century to the end of the 4th century BCE. On the basis of this experience, the philosopher Aristotle built his concept of democracy, which included the possibility of "ruling and being ruled by turns." In Book IV, Chapter 15, of his *Politics*, Aristotle introduced a sharp distinction between the appointment of magistrates by lot, which was thought to be democratic, and their election, which was considered oligarchic. By democracy, he meant government by all the people, while election entailed government by an aristocracy consisting of the few best, which could lead to a perverse form of oligarchy. This classical criterion was taken up in the 18th century by the French provincial Charles-Louis de Secondat, baron de la Brède et de Montesquieu, who regarded suffrage by lot as "natural to democracy." According to Montesquieu, the advantages of making choices by lot are, first, that it "is unfair to nobody, and [second, that] it leaves each citizen a reasonable hope of serving his country" (*The Spirit of Laws*, Book 2, chap. 2). Likewise, in Book IV of *The Social Contract*, the Genevan philosopher Jean-Jacques Rousseau highlighted the role of lots in an ideal democracy, in which, according to his Athens-inspired, assembly-based model, public offices should be considered "a burdensome charge," and administrative acts should be reduced as much as possible.

The choosing of public officers by lots may have, thus, two types of advantages. First, by having this done frequently and establishing short terms of office, it can produce a high rotation of members in administrative or arbitral posts, thus preventing the formation and self-reproduction of permanent and closed elite, whether in the form of an economic oligarchy or a class of professional politicians. By replacing public officials frequently and opening public jobs to wide layers of society, no one can be blamed for making or implementing unpleasant decisions, but no one can be praised either. Only the assembly members remain ultimately responsible for the consequences of collective decisions.

The second type of advantage to choosing by lots and the subsequent rotation in public offices is the production of wide dispersal of knowledge of political and administrative affairs among the citizens. The experience of learning and becoming familiar with the problems of satisfying collective common interests can be a good platform for further occasions of participation in voting and elections, such as the assembly's decision making and the choice of some other public officials.

Thus, a lottery can be an appropriate procedure to select public officers where there is a clear identification of the common interest of the members of the community, there are relatively low technical requirements to fill some public jobs, and there are alternative solid instruments, such as the assembly of members, through which the community can make other important decisions, including control of those appointed by lots.

In the Athenian democracy, in order to preserve the central role of the assembly to make decisions by acclamation or assent, the Council of 500 members in charge of setting the agenda for the assembly was formed of 50 members selected by lot by each of the 10 tribes, which had evolved from military into basic administrative units. The permanent committee of the Council and its president were also selected by lots. About 600 of the other 700 public officers were also selected by lots from among candidates previously presented, including the following: the 10 members of the archonship, approximately equivalent to the post of the modern attorney general, as well as the body in charge of organizing religious ceremonies, who were appointed by lots from a pool of candidates previously selected by each tribe, also by lots; the tribunal members, chosen by lots from a pool of all adult citizens, who were in charge of passing judgment on the legality of the conduct of public officials; and those holding a number of administrative jobs, encompassing treasurers, those in charge of settling public contracts and collecting public revenues, and those supervising streets or inspecting markets. The procedure of selecting candidates for public offices by lots was based, initially, on candidates drawing white and black beans from a container with an open top. In a further development, Athenians also used allotment machines, usually a tube in which balls could be inserted at random and released at the other end.

Among further occurrences, the first apostles of Jesus drew lots to select the replacement for the traitor Judas Iscariot, according to the Acts of the Apostles. On the basis of this precedent, some early nonorthodox Gnostic Christians drew lots at each of their meetings to elect priests, bishops, and other officers. This device could also be aimed at preventing a sacerdotal oligarchy from developing. But the Christian Church condemned such a practice as blasphemy and solemnly forbade the choice of priests, bishops, or other prelates by lot, more formally after the 13th century.

A number of late-medieval and renaissance city-republics and communes around the Mediterranean Sea used lots for choosing magistrates and allocating officers in charge of implementing assembly decisions. This was, in particular, the case in Venice for the indirect election of the Duke (*Doge*) from the 13th century following direct election by the people's assembly from the end of the 7th century. The popularly elected Great Council adopted an increasingly complicated, 5-day-long procedure to choose the Duke with up to nine stages of approval ballots and lots, which was conceived with the aim of making manipulative maneuvers impossible. Likewise, the Florentine republics during the 14th and 15th centuries, and again in the late 15th and early 16th centuries, elected their main governmental body, the Lordship (*Signoria*), chaired by the standard bearer of justice (*Gonfaloniere*), by means of a complex system of approval ballots and very frequent lots. Again, the aim was to prevent fraud, manipulation, and the commune's domination by a few powerful families. Also, in Barcelona, at least from the 15th century, the popularly elected

Council of 100 chose the members of the Consulate of the Sea, the judicial body for commercial and maritime affairs, by an indirect procedure involving lots.

Finally, lots were still being used, in combination with several stages of indirect elections, in Spain and the Spanish colonies in the Americas in the early 19th century. First, elections called by the central junta formed to organize the resistance against Napoleon's troops were held in 1809, with municipalities electing candidates for deputies that were finally selected by lot. New elections in 1810 to form an extraordinary assembly ("Cortes"), which gathered in Cádiz and produced a new constitution, also involved a combination of indirect elections in three stages and a final selection of one deputy in each district by lots among the two or three candidates previously chosen. After the approval of the so-called constitution of Cádiz in 1812, this type of procedure was not used again in Spain, but it was followed in some further elections in Spanish America. Specifically, in Buenos Aires, indirect elections of colleges (usually called "juntas") led to the final selection of members of the provincial assembly by lots in 1811, while a mixed procedure of voting, lots, and final popular vote by plurality was used for the election of governors in 1815. In Mexico, local elections in 1812 involved some stage of selection of candidates by lots. In Chile, a combination of lots and plurality voting was still being used in 1822. Lots remained the usual practice in indigenous communities that were not politically integrated into the new independent states' political institutions. Ironically, they became part of the supposedly traditional "usages and customs" of the indigenous people to be preserved in the 21st century—although the use of lots was actually the most visible legacy of Spanish colonial rule.

The choice of public officers by lots was replaced by election of representatives based on popular votes as new, increasingly large communities and modern states addressed collective issues of higher complexity and as different interests and values developed among their citizens. In current times, lots are used as a method to distribute goods and responsibilities in some private corporations as well as for allocating temporary jobs, vacation periods, or household tasks in other private settings. For public affairs, they are used in some countries for certain sectoral or relatively minor tasks, such as selecting jurors for jury trials, appointing election administrators, breaking election ties, or selecting candidates for military service.

At higher political levels, similar procedures are used in certain institutional settings in confederal or international organizations when members require that decisions be made only by near unanimity or a broad consensus. Some of them use, in particular, procedures of rotation by turns of high public offices, which, a priori and in the long term, produce the same effect of random selection as lotteries. A major example is the Helvetic Confederation of Switzerland, which is still mainly an instrument for preserving popular self-government of the communes and the cantons, where the presidency of the Federal Council is filled in rotation among the Council members by turns. Likewise, for a long period, the chairmanship of the European Council was held in 6-month turns by the member states. The United Nations Organization also distributes some high offices by informal rotation and turns among its member states. The presidency of the General Assembly is filled by "symmetric rotation" among countries of the five regional groups (Africa, Asia, Latin America and the Caribbean, Western Europe and other developed countries, and Eastern Europe). The Security Council, which works by near-unanimity decisions of the five permanent members, is also formed by a number of temporary members rotating in post for periods of 2 years. They are formally elected by the Assembly, but they must also be distributed fairly from among the different world regions, including two for Latin America and the Caribbean, two for Western Europe, one for Eastern Europe, and five for Africa and Asia. Similar proportions are used for filling, by informal turns, the posts in the Economic and Social Council and other committees.

Josep M. Colomer
Institute for Economic Analysis,
Higher Council for Scientific Research
Barcelona, Spain

See also Consensus; Democracy, Direct; Electoral Systems

Further Readings

Burnheim, J. (1981, Spring). Statistical representation. *Radical Philosophy, 27,* 5–12.

Goodwin, B. (1992). *Justice by lottery*. London: Harvester-Wheatsheaf.
Headlam-Morley, J. W. (2009). *Election by lot at Athens*. Cambridge, UK: Cambridge University Press. (Original work published 1890)
Manin, B. (1997). *The principles of representative government*. Cambridge, UK: Cambridge University Press.
Mulgan, R. G. (1984). Lots as a democratic device of selection. *Review of Politics, 46*(4), 539–560.

Election Observation

Election observation entails "the purposeful gathering of information regarding an electoral process, and the making of informed judgments on the conduct of such a process on the basis of the information collected, by persons who are not inherently authorized to intervene in the process" (UN Code of Practice quoted in Harris, 1997, p. 27).

Thus, election observation is not only central to but also has become part of the electoral process worldwide, especially in emerging democracies. The purposes of electoral observation are to promote transparency and encourage free and fair elections. Openness is an essential element of ensuring confidence in the election process. Election observation helps strengthen the democratic process and institutions, enhances the value of elections, and instills confidence in the outcome of the election (John Dugard, 1998). Election observation is expected to be nonpartisan, so that all interested parties can accept the election observers' findings on the observed elections. For an election observation to thrive, it has to be credible in the eyes of contesting parties. This trustworthiness comes from the election observers' refusal to take sides and the impartial and fair-minded nature of the observation endeavor.

Transparency in the electoral process is of critical importance in ensuring public confidence in the election system and recognition of the election outcome (Carl Dundas, 1994). Election observers might persuade some citizens and politicians to take part in the electoral process rather than to resort to violence, as was the case in Angola and Mozambique. In Zambia, in 1991, observers calmed intense disagreements over electoral rules and regulations by asking for concessions from government and electoral officials on areas of disagreement (National Democratic Institute for International Affairs, 1992). Thus, as Gerhard Totemeyer and Denis Kadima (2000) note, election observation can perform a constructive function in generating an environment favorable for everyone to take part in general elections. Election observation adds value to the electoral process and thereby gives it legitimacy. Moreover, the involvement of observers may also encourage a government that has lost power to recognize the outcome and therefore step down. Nicaragua in 1990 under President Daniel Ortega and Zambia in 1991 under President Kenneth Kaunda are cases in point. In this way, election observation helps promote stability. International observers may be asked to arbitrate disagreements among contesting political organizations in an attempt to lessen hostilities prior to, during, and in the postelection periods. According to the National Democratic Institute for International Affairs (1992), for an election outcome to be acceptable to all interested parties, it should not only be seen as free and fair but should also not be marred by allegations of fraud. Thus, although free and fair elections are not a sufficient condition toward democratic consolidation, they are central to democracy development because of "their ability to jump-start the process of democratisation and boost the morale of prodemocracy forces" (Neil Nevitte & Santiago Canton, 1997, p. 51).

Historical Origins

Previously, journalists, academics, and embassy staff observed elections in foreign countries. However, following World War I, political participation in government came to be accepted as a fundamental right of all citizens; since then, election observation has been institutionalized internationally. The United Nations initially took part in election observation in South Korea in 1948, because monitoring the elections in countries coming out of dictatorial military regimes or authoritarian rules was considered essential. Since then, election observation has become common and is used in developed and developing countries. Subsequently, a number of organizations, such as the Commonwealth in 1971 and the Conference on Security and Cooperation in Europe in 1990, became involved in election observation. There has been an increasing acceptance of the contribution observation makes to the election process not withstanding its limitations.

As M. Kupe (2000) has noted, participation of observers could help decrease propensities toward improper practices by overzealous political parties or government administrators in charge of elections. The involvement of observers during voting and at the count can help calm down volatile election environments and is thought to encourage the openness of both the polling and counting of the ballot papers (Dundas, 1994). The presence of international observers is intended to promote confidence in the process, to discourage electoral irregularities, and to give an account on the impartiality of the elections to the international community. Election observation ensures that the electoral outcome is credible internationally in the light of specific democratic conditions imposed by donors who provide economic assistance and other help. This explains why election observation has become so important in recent years, especially in countries undergoing political change.

In established democracies, election observation seeks to act as an example to emerging democracies. Election observation is not merely about verifying the fairness of an election; instead, and more crucially, observation efforts strengthen the legitimacy of democratic norms and procedures to thousands of people and, thus, help in building and strengthening a culture of democracy. It is in this context that election observation has come to be accepted as a key part of the democratic process because such efforts seek to promote transparency and accountability—the central tenets of democracy.

Election observation missions take different shapes. Their diversity depends on the type and size of the organization involved, the organization's period of stay in the country holding elections, and the kind of report the organization produces after the elections. Nevertheless, there are certain experienced and credible election observers—the Commonwealth, the European Union, and the Carter Center for Democracy in Atlanta, Georgia—that have greater resources for observation than others. Generally, election observers oversee the administrative arrangements; preparations of the electoral authority; the behavior of election officials, police, party agents, and voters; the sealing of ballot boxes before and after voting; the escorting of boxes to the counting center; and counting (The Catholic Justice and Peace Commission, 1994). The observers are expected to produce a report, which highlights the problems noted, and make recommendations. They are also expected to determine whether the elections were free and fair. As Kupe (2000) states,

> Election observers are "watchdogs working for the electorate and the political parties involved in the elections. It is their business to make sure that elections are conducted properly during the prescribed times and at the designated venues. It is the observer who reports directly or indirectly to the outside world regarding the fairness, or otherwise, of the election process. (p. 102)

Types of Election Observers

There are two main types of observer groups: local and international groups. Local observers, who are normally citizens of the country whose elections are being monitored, have an advantage over international observers in that they are familiar with the country concerned. However, although local observers promote transparency in the electoral process, they are often treated with suspicion as having a different agenda and thus not seen as impartial. On the other hand, international observers "have added a new dimension to election transparency" (Dundas, 1994, p. 45). Nevitte and Canton (1997) observed that "public confidence in internationally driven [observation] efforts characteristically hinges on the reputation and legitimacy of the international or regional organization involved, and derives in large part from the multinational membership of the observation team in place" (p. 50). For Phiroshaw Camay and Anne Gordon (1999) "international observers bring an added credibility to the monitoring and assessment of elections, in that they are able to refer to their experience elsewhere and apply international standards of good practice wherever they go" (p. 259). Such observations demonstrate that international observers act as a stamp of approval.

Moreover, observer groups can use different approaches to observe elections. These can take the form of regular short visits, permanent groups, mobile teams, or stationary teams. Nevertheless, election observers face a number of limitations. First, as Dundas (1994) noted, "there are limitations to the extent to which observers generally can impact on the transparency of the system, since they are not in charge of the machinery which runs the election." Second, the other limitation faced by election observers is that "they have a limited time to see only the final days of the campaign leading up to the polls, and sometimes many leave before the final results are known" (p. 45). This limit on the time available for observation compromises the role of election observers in promoting transparency and accountability in the electoral process.

Election observation is common in countries undergoing political transformation from autocratic regimes to multiparty systems. The uncertainty and confusion in these countries create concerns about whether election administrations can deliver accurate and impartial electoral results. International observation of elections in Africa initially occurred as nations made the transition from colonial rule to independence (in Zimbabwe in 1980, Namibia in 1989) or when there was no centrally controlled authority, as in Uganda in 1980. In 1991, Zambia became the first independent African country to ask for international election observers. According to the National Democratic Institute for International Affairs, local and external election observers can help ensure that elections are free and fair, which in turn can contribute to the development of accountable, effective governance. International election observation provides evidence of the extent to which a regime is committed to a variety of democratic values and procedural norms. It also gives voters a critical opportunity to challenge authoritarian governments (Nevitte & Canton, 1997). Such observation occurs primarily in cases where there are doubts that a free and fair election can be conducted (Ananias Elago, 1999). As Bojosi Otlhogile (1994) notes, these transitional countries are attempting to reconstruct themselves politically and seek affirmation from observers that elections were fairly conducted and that they meet the minimum requirements of an international standard. Such validation can help these countries regain their places in the international community.

As Otlhogile notes, countries undergoing transformation attract international observers because there is unease surrounding an election in such a country. Observation is meant to reestablish confidence among those contesting such an election as well as among foreign investors and the international community generally.

Notwithstanding the importance of election observation in promoting transparency and accountability in the electoral process, a few observations can be made about election observers. Although election observers are widely appreciated and at times overemphasized, their findings have never resulted in a reelection in any country, regardless of the assessment. The April and December 2003 elections in Nigeria and Russia, respectively, are cases in point. In fact, election observation has become a standard institution that is expected to contribute to the democratic process, especially for the reasons discussed above. Moreover, even if they are expected to be impartial, an element of bias cannot be ruled out. However, their reports might indirectly prompt people to revolt, resulting in a new election.

David Sebudubudu
University of Botswana
Gaborone, Botswana

See also Democratization; Election Research; Electoral Systems; Transition

Further Readings

Camay, P., & Gordon, J. A. (1999). *The people have spoken . . .: A review of the 1999 South African election*. Johannesburg, South Africa: Co-operative for Research and Education.

Dugard, J. (1998). Current issues in election management. In C. W. Dundas (Ed.), *Discussion of election issues in commonwealth Africa*. London: Commonwealth Secretariat.

Dundas, C. W. (1994). *Dimensions of free and fair elections*. London: Commonwealth Secretariat.

Kupe, M. (2000). Does election observation improve the confidence of political parties and the public in election results? In C. W. Dundas (Ed.), *Rules of elections in Commonwealth Africa*. London: Commonwealth Secretariat.

National Democratic Institute for International Affairs. (1992). *The October 31, 1991 national elections in Zambia*. Atlanta, GA: Carter Center, Emory University.

Nevitte, N., & Canton, S. A. (1997). The role of domestic observers. *Journal of Democracy, 8*(3), 47–61.

Otlhogile, B. (1994). Observing for democracy: A note on the practices of the commonwealth observer groups. *African Journal of International and Comparative Law, 6*(2), 293–296.

Totemeyer, G., & Kadima, D. (2000). *Handbook for election observer missions.* Auckland Park, South Africa: Electoral Institute of Southern Africa.

Election Research

As the defining institution of democracy, elections have always been a major focus of attention both for political theorists and analysts—and for journalists and practicing politicians. Indeed, it is hard to talk about democracy, either generally or at a national level, without referring to elections. This entry focuses on specialized research, mostly quantitative, into voting at general elections. It thus passes over the interesting and growing literature on direct policy elections (referendums and initiatives) since they are still not a major feature of contemporary democracies except in Switzerland and, at the state level, in the United States. On the other hand, the choices voters make are framed by political parties, so the entry has to consider their activities, particularly as they impinge on voting.

Demographics and Political Geography

While analyses of party organization and campaigning appeared from the latter part of the 19th century onward, the tradition of systematic quantitative studies of voting statistics began with André Siegfried in France, Harold Gosnell in the American Midwest, and V. O. Key in the American South from 1900 onward. Such analyses used demographic data—social and voting statistics, generally from the smallest constituency unit available—to make sophisticated inferences about individual voting behavior and its influence on both personal characteristics and political history.

This tradition of research, however, passed out of the mainstream under the impact of two developments. The first was the discovery of the ecological fallacy involved in inferring individual behavior from aggregate statistics. For example, the finding that constituencies in the United Kingdom with a large Black population voted disproportionately for the racist British National Party (BNP) does not imply that the BNP attracts non-White votes. On the contrary, it shows that Whites in such constituencies vote for it disproportionately, under the impact of high immigration. It is hard to tell from aggregate statistics what exactly is going on at the individual voter level. The second development relates to the use of mass survey techniques.

Survey-Based Research and Social Group Theory

In addition to difficulties involving the ecological fallacy, the problem for the statistical foundations of demographic research was compounded by the application of mass survey techniques to study individual voters directly. This approach was pioneered by Paul Lazarsfeld and his associates in Erie County, New York, for the U.S. presidential election of 1940, and replicated in Elmira, New York, in 1948. Not only did these studies demonstrate the feasibility of applying statistical techniques of sampling and quantitative analysis to voting, but they also demonstrated the way in which data collection and analysis had to be driven by theory to produce relevant conclusions or even to develop a questionnaire in the first place.

Coming from a market research background, Lazarsfeld envisaged voters as behaving like consumers choosing between tins of beans on the supermarket shelves. Relying solely on advertising for their information, brands promoted more in the media would sell better than ones promoted less. The 1940 questionnaire thus featured many questions on voter media exposure (with parallel content analyses of actual media coverage of candidates), with a few questions on social characteristics thrown in for classification purposes. To their surprise, the research team found few and weak correlations between media coverage and voting behavior but strong correlations with class, religion, and urban versus rural residence.

In response, they jettisoned the media and adopted instead a social group theory of voting behavior, measured through an Index of Political Predisposition (IPP), which very roughly measured probabilities for individuals with varying combinations of social characteristics voting Democrat or Republican. On this basis, the questionnaire for the 1948 U.S. presidential election sought to understand how voters' flow of information and voting preferences were structured by their social group membership, pulling them back to the historical group choice in the course of the campaign. Social group theory reached its apotheosis, with the help of computers, when Ithiel de Sola Pool, Samuel Popkin, and Paul Abelson pooled all existing public opinion surveys of American voters to create demographic profiles of the U.S. states and the impact of current issues on small subgroups of voters. On the basis of these simulations, they correctly predicted the result of the 1960 and 1964 presidential elections in advance, achieving high correlations with the actual electoral college votes of the non-Southern states.

The Michigan School: Party Identification and Issue Impacts

The publication by Angus Campbell, Philip Converse, Warren Miller, and Donald Stokes of *The American Voter* in 1960 not only substituted national mass surveys for local ones but also proposed an alternative conceptualization of voting choice to social group theory. It was not that this "Michigan School" ignored the influence of social factors. They saw them, however, as lying rather far back in the "funnel of causality" preceding the voting act. Their influence was channeled through an individual's party identification—their affective orientation to an important group object in their environment. Social influences were thus transmuted into individual psychology. Party loyalty then colored the voters' reactions to immediate influences on how they viewed the candidates and issues of the campaign. The change in these from the end of the 1940s to the 1950s explained the landslide election of the Republican Dwight Eisenhower after the Democratic rally of Harry Truman in 1948.

This conceptualization of voting as essentially individual and psychological in nature dominated empirical research on elections in both the United States and the rest of the world for the next 30 years. In a modified form, it continues to this day. Its dominance is only partly explained by its comprehensive reach and intellectual power. It was also embedded in the design of questionnaires for the mass election surveys now carried out regularly in almost all the developed democracies of the world. The Michigan authors themselves were active in stimulating and organizing such surveys in Europe, from whence they spread to Asia, Latin America, and Australasia.

For the first time, such surveys provided an insight into the political thinking and behavior of ordinary individuals, complemented by the massive increase in both the number and sophistication of public opinion polls. In addition, for the first time, successive surveys provided information about the dynamics of opinion change between elections and the effect this had on party support. Such over-time analyses required continuity in the questions asked at each election, simply to provide comparable information over long periods. This embedded the Michigan School's conceptualizations and measures even in studies that did not necessarily share their preconceptions or that even set out to challenge them.

Party Identification and Voting Predispositions

The major challenge was to understand the concept of party identification and whether it really differed, methodologically or conceptually, from voting intention. Michigan-style surveys carried out in Europe seemed to indicate that the two were correlated and fluctuated together. If so, what was the point of distinguishing the two?

One answer was to generalize from the specific concept of party identification to the idea of enduring predispositions toward voting in a certain way. The majority of electors vote in the same way from election to election. At the same time, some change their vote, providing the dynamics for election change and party competition. It therefore seems that both predispositions and specific campaign influences such as issues and candidates have to be brought into the explanation of voting choice. A group of largely European scholars in *Party Identification and Beyond* (2009) argued that both Campbell and his associates and Lazarsfeld and his associates had been trying to do just that. Both party identification (PI) and the IPP were trying to measure predispositions, while both allowed for the differentiated impacts on strong partisans and weaker ones of short-term campaign effects or election cues.

Issue Voting and Rational Choice Theory

This recognition of short-term effects rendered the debate of the 1970s on the "replacement" of PI by "issue effects" from new issues, such as Vietnam, gender bias, and the environment, rather superfluous. It did, however, mark the emergence for the first time of purely theoretical influences on voting research as distinct from the theory-driven data analysis that had dominated the field up to then. These influences stemmed from the seminal presentation by Anthony Downs of *An Economic Theory of Democracy* (1957)—a collection of many theories and models, not generally reconcilable with each other, of how office-seeking parties competed for votes and electors cast votes for the party closest to them on policy.

The most influential of Downs's models (because it seemed to fit Anglo-American politics so well) was where two parties competed with each other on policy in a one-dimensional left–right space, along which electors ranged themselves in a symmetrical unimodal distribution, with the central peak coinciding with the median (middle) voter position. To form a majority of 50% plus one, the median vote was crucial, so both parties would move toward the more moderate position of the median voter to secure a majority. This accounted for the consensual bargaining and policy compromises of the Anglo-American democracies, as contrasted with the *immobilisme* and ideological rigidity of multiparty governments.

Turnout

Downs's two-party spatial model had an increasing influence on election research from the 1970s onward, in line with the growing dominance within political science of rational choice reasoning and mathematical modeling. One aspect of his ideas was not closely pursued, however—his demonstration that it was not rational to vote. He assumed that electors wish to consider their policy preference and decide whether to vote on the basis of their utility from this, discounted by the probability of their vote changing the election outcome. Since this probability is always vanishingly small, nobody should vote. As the majority do (in general elections at least), other factors, principally a sense of duty, have to be drafted in as explanations—supported so far as they go by the survey data.

Economic Voting

Rational choice theorists seemed to have little difficulty with incorporating duty or altruism into their equations on turnout. On the other hand, voting choice seemed potentially explainable by self-interested, primarily economic, calculations. Government's ability to sustain prosperity and growth, increase employment, and avoid inflation were increasingly seen as the main drivers of vote. Conceptually, theorists also seemed capable of linking political science more closely with economics in the study of political economy.

While such analyses could be and were carried through at the individual level, on survey data both economic indicators and voting statistics were readily available in aggregated form. Where necessary, survey responses themselves could be aggregated to give preference and policy time series. As a result, the focus of empirical research shifted—from surveys and individual-level analysis to the aggregate level—in the 1980s. As the major objective was to explain (and preferably predict) the overall election outcome, aggregate equations and econometric methods formed a more direct way to approach this. Moreover, the general availability of aggregate indicators lent a comparative dimension to research—all governments published them, whereas election surveys were not always carried out in the countries of interest.

Despite 2 decades of intensive analysis, however, little progress was made toward the ultimate goal of a unified theory of economic voting. A review of around 2,000 concluded that economic conditions produced contradictory and conflicting political results. The main generalization to emerge from numerous studies over different countries was that governments generally lost votes at the rate of 2.2% of their previous vote. The loss was the cost of governing, expressed in economic fashion but hardly due to exclusively economic factors.

Predispositions and Cues at the Aggregate Level

One reaction to the failure of economic trends to produce consistent voting responses was to regard them as one type of issue among others. A comprehensive classification of issues could be made on the basis of surveys and newspaper reports, and a universal direction and weight can be assigned to each issue type. By seeing what issues were prominent in each campaign, the result could be predicted and compared with the actual outcome.

The idea that certain issues always favor certain parties stemmed from a salience theory of party competition and elections, whereby parties always strove to make their own issues prominent and downgrade those of their opponents. However, the number of votes affected by issues prominent in the campaign was always less than the basic vote of the party—the average vote it received independently of issue effects. This vote was explained as being cast by core supporters or party loyalists—party identifiers or core social group voters under another name. In this way, the old survey-based frameworks reemerged at aggregate level under other names—still necessary concepts, however, to cope with the stability and change simultaneously present in election results.

Ian Budge and Dennis Farlie's attempt to explain and predict elections in this way remained an isolated venture. They used the same logic in ad hoc models designed to predict the outcome of a general election. In Britain, these took the form of the Essex model. This was a multivariate regression equation, which allowed previous voting support, economic expectations, and tax levels to explain all the month-to-month fluctuation in voting intentions that they could. Sharp falls or increases in voting support were identified with events that occurred at that time (e.g., the Blair effect following Tony Blair's election as Labor leader in 1994). The combined equation, modified for each election, was used to predict votes in the 1997 and 2001 general elections successfully. This predictive success was bought at some cost in terms of generalizability and explanatory power (why should Blair *have* this effect?). However, the overall approach to prediction is itself generalizable to other contexts.

The Macro Polity: An Update of the American Voter at the Aggregate Level

The same approach involving econometric modeling and aggregate data was used in adapting the concerns of *The American Voter* to modern times. Using basically the same framework as *The American Voter*, *The Macro Polity* (2002), by Robert Erikson, Michael McKuen, and James Stimson, models voting at the aggregate level as the effect of aggregate predispositions to vote for a party and the effects of issues, particularly economic issues, in changing votes. Reflecting a growing recognition that voting has to be put in context, *Macro Polity* also concerns itself with the end result of the process—how election outcomes affect government policy intentions and enacted policy—and even with how these match up with voter preferences, thus addressing the central concern of democracy. A major technical achievement of the project has been to devise a measure of the (left–right) policy mood of the electorate from hundreds of different policy questions asked in opinion polls over the past 50 years. This not only provides a basis for explaining voting behavior but also for investigating the extent to which public policy reflects popular preferences.

The limitations of *The Macro Policy* are also those of its predecessor—it focused on the politics of one country over a limited period of time. This limits the number of cases to 12 (postwar presidential elections). It also runs the danger of tying its investigation too much to the specifics of U.S. politics, which, as *The American Voter* showed, have their peculiarities. Parallel studies are currently under way in Britain and France. One may hope after 40 years that they do not stir up the same controversy over concepts and measures as party identification did. So far, policy mood seems to apply as well in Europe as in the United States.

Multimotivated Voting

When reading through long lists of detailed and variable findings from the election studies, one is sometimes tempted to ask just what the detailed and expensive analyses of elections over the past 5 decades have shown. There is much confusion about this. One consistent lesson, however, from the earliest to the latest studies is that voters are multimotivated. That is, they do not vote on their policy preferences alone, and some votes in fact may be cast on grounds unrelated to policy. This is not solely due to electors themselves having nonpolicy concerns. To a considerable extent, the institutional structure of general elections requires them to settle nonpolicy questions: Who is best at governing us? Which party has the most appealing and trustworthy candidate? Which has best ridden out the most recent scandal? All these factors may enter quite legitimately into the voting decision.

The fact that voters are multimotivated puts current rational choice models largely out of court so far as explaining elections are concerned. They all postulate that voters are purely policy orientated. If they are not, parties have no need to converge, and the median preference (or *seeming* median preference, as the declared distribution of votes cannot be unambiguously interpreted as a policy distribution) loses its preeminent position.

Representational Consequences

Elections are at the heart of the democratic representational process and are thus expected to reveal and enforce popular preferences. As Bingham Powell (2000) notes, recognition of their role has prompted the broadening of election research noted above, to the extent that elections are increasingly being studied in their full representational context as "instruments of democracy."

According to Joseph Schumpeter (1942), here the recognition of voters as multimotivated and not driven by pure policy concerns poses an acute dilemma for representative democracy. If elections do not clearly reveal the popular will so far as public policy is concerned, how can the democracy be directed to effect it? Or do we simply define democracy as a struggle for power between competing teams of leaders decided by voters in elections—who have, however, to put up with whatever the policy leaders decide on subsequently?

The solution may be to recognize, with Powell, that elections express the popular will in two ways:

1. the *proportional* way, operationalized in the preference of the median voter (even though the median position, which emerges from the declared distribution of party votes, may be a bit unstable and subject to challenge), and
2. the *majoritarian* way, in which the popular will is identified with the policy preference of plurality party voters.

The plurality voter position may also be subject to challenge since nobody knows how many votes cast for the largest party were actually cast on nonpolicy grounds. In the end, however, the declared distribution of votes has to be accepted as providing authoritative guidance on policy preferences as on other matters as well. With no way of deciding between median and plurality claims to represent the popular will, researchers have to accept both and arrange compromises between them. How this is done, however, is more a matter for research into governments than into elections.

Election Data

While election research has generated and tested many theories of the determinants of voting choice, its most significant achievement to date could well be described as the generation of vast amounts of information, primarily survey data, but broadening out to time series and comparative collections of economic indicators, voting statistics, government expenditures and policy indices, personnel and structures of governments, text-based counts of policy emphases, campaign issues, manifestos, and speeches—the range and quantity of the information is truly staggering and still underexplored. The combination of information at the several levels of electors, parties, governments, and policies gives political scientists the ability to settle many of the unanswered questions raised in the previous discussion. If there is not yet a complete theory of democratic decision making and of the role of elections within it, researchers certainly now have the ability to formulate and check one against the material now available—not just for any one country but for the whole range of democracies. This is perhaps the major product of election research over the past 50 years, which has contributed cumulatively to the understanding of voting choice and of the democratic electorate.

Ian Budge
University of Essex
Colchester, United Kingdom

See also Party Manifesto; Political Psychology; Rational Choice; Representation

Further Readings

Berelson, B. R., Lazarsfeld, P. F., & McPhee, W. N. (1954). *Voting*. Chicago: University of Chicago Press.

Budge, I., Crewe, I., & Farlie, D. J. (Eds.). (2009). *Party identification and beyond*. Colchester, UK: ECPR Press. (Original work published 1976)

Budge, I., & Farlie, D. J. (1983). *Explaining and predicting elections*. London: Allen & Unwin.

Campbell, A., Converse, P. E., Miller, W., & Stokes, D. E. (1960). *The American voter*. New York: Wiley.

Downs, A. (1957). *An economic theory of democracy*. New York: Harper.

Erikson, R., McKuen, M., & Stimson, J. (2002). *The macro polity*. Cambridge, UK: Cambridge University Press.

Pool, I. de S., Popkin, S. P., & Abelson, P. (1965). *Candidates, issues and strategies*. Cambridge: MIT Press.

Powell, G. B. (2000). *Elections as instruments of democracy*. New Haven, CT: Yale University Press.

Schumpeter, J. A. (1950). *Capitalism, socialism, democracy*. New York: Harper.

Elections, Primary

Primary elections—or preliminary, earlier elections—can come first both in order and in importance. They come first in order because selecting candidates is one of the first things that political parties must do before an election. They can be first in importance as well because those who are eventually elected to office are the successful candidates whom the parties previously selected; and they are the ones who will determine what the party, the parliament, and the country's politics will look like until the next elections.

In the study of preliminary elections, the unit of analysis is the *single* party in a particular country at a specific time. There are very few established democracies where the legal system specifies criteria for candidate selection—for example, Finland, Germany, New Zealand, and Norway (until 2002)—and only in the United States does the legal system extensively regulate the primary elections. Preliminary elections, or candidate selection methods, are thus the nonstandardized and predominantly unregimented party mechanisms by which political parties choose their candidates for general elections. The result of this process is the designation of a candidate, or list of candidates, as the candidate(s) of the party. The party then becomes effectively committed to the candidate(s) and to mobilizing its strength behind the chosen candidate(s).

This entry delineates the most important distinguishing variable for explaining how political parties choose their candidates—the selectorate, that is, who is allowed to take part in the selection of the candidates. Along with another important variable, that of candidacy requirements, these two criteria are akin to the supply and demand sides of candidate selection. While candidacy narrows the supply of contestants who can be selected, it is the selectorate that decides who will eventually face the voters in the general election. Due to the scope of its political consequences, the selectorate is the most significant and far-reaching explanatory variable when it comes to primary elections.

The selectorate continuum starts with the most inclusive body—all voters—and ends with the most exclusive one—the single leader. Each part of the continuum, and the variations along it, is illustrated by empirical examples from both established and new democracies. This entry begins by focusing on the well-known case of primary elections in the United States, which reflect the most inclusive of selectorates. It then expands to cover other democracies and their use of preliminary elections. The entry concludes with a short analysis of the political consequences of different selectorates.

Classifying Primary Elections

Primary elections can be classified by distinguishing the five archetypical kinds of selectorates, from the most inclusive to the most exclusive:

Figure 1 Party Selectorates in Primary Elections

1. *The most inclusive selectorate: voters.* This selectorate includes the entire electorate that has the right to vote in the general elections.
2. *The highly inclusive selectorate: party members.* This refers to dues-paying party membership in its European sense, implemented and controlled by the party itself, rather than simply registration as a party supporter as in the United States, which belongs to the above category.
3. *The in-between selectorate: party delegates.* This selectorate is composed of representatives selected by the party members. They can be members of party agencies (e.g., conventions, central committees, congresses) or delegate bodies that were selected for this purpose alone.
4. *The highly exclusive selectorate: the party elite.* This includes small party agencies and committees that were indirectly selected or whose composition was ratified by wider party agencies and other less formal groups.
5. *The most exclusive selectorate: a nominating entity of a single leader.*

Voters

The primary elections in the 50 states of the United States provide most of the examples on the inclusive end of the continuum. The exact location of American primaries depends on the restrictions defined by the different state laws. At the extreme end are the American nonparty primary elections, used in Louisiana from 1978 to select candidates for Congress. In these primary elections, every registered voter could vote for candidates from any party. Blanket primary elections—used in Washington (since 1938), Alaska (since 1968), and California (1998, 2000)—are also at the extreme inclusiveness pole. Here, voters receive a single ballot listing all the candidates from all the parties and decide, for each post separately, which party candidate to vote for. In both the nonparty and blanket primaries, participants do not need to declare their party affiliation in order to take part in candidate selection.

Open primary elections—used in Hawaii, Michigan, Vermont, and other states—are slightly less inclusive than the two previous kinds. As in nonparty and blanket primaries, the voters are allowed to decide in which primaries they want to take part without the need to announce their partisan preference. Yet, unlike these previous types, voters can take part in the primaries of only one party.

Semiclosed primary elections—also used by several states such as Arizona, Indiana, and Rhode Island—require participants to declare their party affiliation only on the selection day and/or allow independents to take part in the party primary they announce their wish to vote in. Semiclosed primary elections move ever so slightly away from the inclusive end of the selectorate continuum because voters need to publicly affiliate with a political party.

The inclusive end of the continuum also includes examples from Iceland, Taiwan, and Spain. From 1971 on, parties in Iceland adopted primary elections in some, and sometimes all, electoral districts, where every citizen in a particular electoral district could participate. In Taiwan, the National Party adopted inclusive primary elections in 1998. The Catalan Socialist Party in Spain opened its candidate selection to “registered sympathizers”—nonmembers who could register as party supporters without paying a membership fee.

American closed primary elections, which demand that voters register according to their party affiliation before the day of the primaries, are located further away from the inclusive end. The level of inclusiveness of the selectorate of the Democratic Progressive Party in Taiwan (1998–2001) also places it at this point. It used a unique weighted selectorate that combined the results of both a vote of the party members and a public opinion poll.

The open convention, which allows any voter to take part in a candidate selection meeting, is probably a less inclusive version of the most inclusive family of selectorates. While it allows any voter to take part in a selection meeting without the need to prove party affiliation or even to preregister, it is still quite a demanding system as it requires the voter to attend a meeting at a certain date, time, and location. This kind of selection—which is similar to the American caucus—was used in the past in Canada, in the 1920s to 1950s period, before the parties began institutionalizing their membership.

Party Members

This discussion now moves into the party members' zone. At its inclusive end, it comes close to including all the voters—as in the case of the Dutch Democrats 66, which allowed party members to vote via postal ballots but then held meetings where both party members and all voters could participate. Another example is the Argentine Peronist and Radical parties, which in some districts (1983–2001) allowed party members and independents to participate in their primaries.

In the middle of the party members' zone is the typical European closed primary elections, which—as opposed to American closed primary elections—usually mean "party primaries" in which the selectors are party members, not just supporters. From this point on in the selectorate continuum, mere party supporters are excluded. The purest type of party primary is where the party members' votes alone determine the candidates. A number of parties have used such party primaries, albeit not consistently over time and not necessarily in all the districts. Among them are the following examples: the Australian Labor Party; the Belgian ECOLO; the German Social Democrats, Christian Democrats, and Greens at the single-member district level; the Icelandic Independence, Social Democratic, Progressive and People's Alliance parties; the Israeli Labor, Likud, and Kadima parties; and the Mexican Institutional Revolutionary and the Democratic Revolution parties.

When party members have a dominant or significant role in candidate selection—but are not the sole selectors and other, more exclusive party actors take part in the selection of candidates—the selectorate remains in the party members' zone but is moving toward the party delegates area. For example, in several of the Danish parties, from the 1970s until recently, central party agencies could veto or change the selection made by the party members. In Finland, the election law states that party organizations have the right to change up to one fourth of the candidates selected by the members' vote. In Canada, national party leaders have some power over candidate selection, although they usually refrain from exercising it, while in Ireland the national party leadership kept and even enhanced its power in those parties that adopted a membership ballot.

Cases where party agencies first filter the candidates who are then put to a membership vote are close to the middle between the party members and the party delegates zones. If the screening process still leaves a large and viable pool of candidates from whom the party members can make the final decision, then the selectorate is still on the party members' side. Since the mid-1990s, the Israeli Meretz party has produced a sizeable "panel" of candidates from which the members choose the final list. The Social Democrats and the Liberals in Britain did much the same, as did several of the parties in Belgium in some districts, particularly the Belgian Socialist Party during the 1960s.

When equal weight is given to the party delegates and to the party members, the selectorate is in the middle between these two categories. In Taiwan, in 1995–1996, the Democratic Progressive Party used a weighted method in which the vote of the members was equally weighed to the vote of the party representatives. British Labour's use of a multistage method since 1997—where candidates are screened by party agencies, selected by party members, and can then be vetoed by the National Executive Committee—can also be seen as a middle-of-the-road example. The same seems to be the case for the Democratic Party of Botswana in 2002, where party members selected candidates after a national party agency screening, and for both the Labor (1960–1964) and the Pacifist Socialist (1957–1973) parties of the Netherlands, which allowed their national executives to propose the list of candidates but let the members then vote and alter both the rank and the composition of these lists.

It should be noted that the party members' selectorate can be further distinguished according to the restrictions on party membership. For example, one rule that could restrict membership—or the right to participate in primary elections—is the rate of membership dues. Members' participation may also be restricted by the requirement of a minimal party membership period prior to candidate selection, proof of party activity, and so on.

Party Delegates

When the party members have less of an impact than selected party delegates, the selectorate is still located between these two zones but is closer to the latter. Here, for example, the members can ratify or reject a list of candidates drawn up by the party agency, as was done by the French Socialist Party in 1986. Since the 1980s, the British Conservatives use a multistage method, which starts with a screening by a nonselected national party agency, followed by a locally selected party agency screening, and finally a party members' selection meeting.

The selectorate moves slightly closer to exclusivity when the candidates are produced by a wide delegate convention, as was the case in Ireland with Fianna Fail, Fine Gael (until it adopted membership ballots in the 1990s), and Labour. Another example is selection by a party agency that might be followed (or preceded) by a membership ballot, as implemented by the Swedish Communist/Left Party.

When the selectorate is an agency of the party, the selectorate is in the middle of the selectorate continuum. Inside the party, the relative size of each agency is a sign of its inclusiveness: Conventions are usually larger than central committees, which in turn are usually larger than executive bodies, such as bureaus. As the size of the particular party agency gets smaller, the selectorate moves toward the exclusive pole of the continuum. The terminology used in each country is not necessarily equivalent, and hence one must be cautious when inferring the extent of inclusiveness based solely on what a particular party calls a specific agency. The use of party delegates is widespread. Since the 1950s, the major German parties have used delegate conventions at the single-member district level. This was also the typical selectorate in Australia. Several Israeli parties use their central committees to select their candidates. This was also one of the methods used by the Canadian parties 60 years ago (1920–1950), which have since adopted the more inclusive membership ballot.

When a nonselected party agency—such as a nominating committee—has an influence on the selection of candidates alongside the selected party agency, the selectorate moves toward more inclusiveness but remains within the party agency zone. One example is the assorted system used by both the National Liberation and the United Social Christian parties in Costa Rica, where most candidates were chosen by a selected party agency, but several were nominated by the party president. There are many examples of both directly and indirectly selected party agencies taking part in candidate selection, as in the Austrian Socialist Party (1945–1990); the Dutch Christian Democrats (1986), Christian Historical Union (1960–1979), and Radical Political Party (1973–1989); and the British Conservatives from the 1950s to the 1970s and British Labour from the 1950s until 1987. Yet another example is the New Zealand Labour Party, which has used a method since the 1950s where both nominated and selected delegates choose the candidates.

The selectorate is between the party delegates and the party elite when there is a relative balance of power between the selected and the nonselected party agencies. This was the case with the methods used in the 1980s in the Union for French Democracy, the Dutch People's Party for Freedom and Democracy from the 1960s to the 1990s, and the Spanish Socialists from 1979 to 1998. The Norwegian parties, in most cases, from the 1920s and until 2002, followed such a system by virtue of a law that allocated funding to political parties that selected their candidates in a (selected) delegate convention at the level of the multimember constituency. Candidate selection at the länder level in Germany also belongs here—although party delegates are the final decision makers at this level, the selection is made on the basis of a recommended list that is designed by the land party elite.

Party Elite

This category includes nominating committees that are formed for the sole purpose of selecting the party's candidates as well as nonselected party agencies (typically small executive boards) that are entrusted with several different tasks, including the selection of candidates. The composition of both nominating committees and nonselected party agencies can be regarded as slightly more inclusive if it is indirectly selected or somewhat more exclusive if it is not.

When the selection power of the party elite is stronger than that of the selected party agency, the selectorate moves into the party elite zone and is within the exclusive part of the continuum. There are cases of methods that involve various selected and nonselected (or highly indirectly selected) party agencies, with more influence given to the latter. For example, the French Rally for the

Republic in the 1980s allowed a nominating committee to select its candidates but afforded some influence to other directly and indirectly selected party agencies. The Italian Communist Party from 1956 to 1986 allowed for the involvement of a selected party agency, but the final word on candidate selection was given to an indirectly selected party agency. The Chilean Party for Democracy and the National Renovation Party allowed their national councils to be involved, but the final decision was made—because of the constraints set by the binominal electoral system—by negotiations between the party leaders. When, for example, the selected party agency is only asked to ratify a decision made by the nominating committee, the selectorate is still leaning slightly toward the inclusive side of the party elite zone. This was the case in the Japanese Liberal Democratic Party in the period between 1950 and 1990.

When only a nonselected party agency/group is involved in candidate selection, then the selectorate is squarely in the middle of the party elite zone. This was the case in the Venezuelan Democratic Action Party in the 1990s; the Italian Christian Democrats from 1957 to 1984; the Mexican Institutional Revolution Party before 2000; the Chilean Independent Democratic Union from 1989 to 2001; in a significant number of constituencies in several of the main Belgian parties, especially since 1968; the Indian Congress Party in the 1950s and 1960s; and the Danish Peoples Party in 1998.

The selectorate moves toward the single-leader pole with such exclusive selectorates as a gathering of the party founders in a new party or an informal group of factional leaders in older parties. Israel's ultrareligious parties serve as an example of such highly exclusive selectorates.

Single Leader

The extreme end of the exclusive pole is defined by a selectorate composed of a single individual. If the leader does not have complete control over candidate selection, then the selectorate is close to the exclusive end of the selectorate continuum but not at its pole. In Forza Italia, in the 1990s, the founding leader Silvio Berlusconi chose the candidates in cooperation with the party's regional coordinators. The French National Front leader, Jean-Marie Le Pen, together with the party's general secretary, chose the party's candidates, with some influence given to nonselected forces. Winston Peters, New Zealand First's founder, was also given almost complete control over candidate selection.

Like the party elite, a single leader will lean more toward the inclusive side if the leader is selected and more toward the exclusive pole if the leader is not a selected leader. In 2005, Israeli Prime Minister Ariel Sharon quit his party (Likud) and formed a new one (Kadima). Since he also quit his position as prime minister and called for early elections, time constraints and a new party allowed him to choose the list of candidates.

Political Consequences

Different levels of selectorate inclusiveness produce different political consequences. For example, in terms of quantity, the more inclusive selectorates are the more participatory ones. In terms of quality, though, the picture becomes less clear because most of those who do join parties do not participate in their party's candidate selection process and are not affiliated for more than a short period.

Smaller, exclusive selectorates are able to balance representation, which is more difficult when the aggregation of a large number of votes produces the party candidates. Party agencies are actually more competitive than primaries because of the shorter "distance" between the candidates and their selectors.

Inclusiveness also influences responsiveness. Legislators selected by a small selectorate owe their positions to that selectorate and are likely to be party players. Legislators selected in primaries need to reach a massive, fluid audience and will behave more like individuals than team players.

Conclusion

The consequences of primary elections can be more or less significant as a result of the characteristics of a particular nation's politics. For example, in countries with single-member districts, if the number of safe seats is either large or growing then the selection process of the winning party could be more decisive than the election itself. In both the United States and Great Britain, more than one half of the constituencies are safe for one party or the other, with majorities of greater than 10%, which means that the effective choice of who will become a legislator is made not by the voters in the general election but by the selectorate in the primary elections. Moreover, even in marginal seats, the correct choice of a candidate by a party could make the decisive difference between winning the seat and losing it. In countries using proportional electoral systems, the selection of candidates at the top of a party list can virtually guarantee election, particularly in the major parties, practically regardless of the results of the general election. In short, in the majority of democratic nations, in a majority of the parties, the preliminary selection inside the party is as important as the subsequent election by the general population.

Reuven Y. Hazan
Hebrew University of Jerusalem
Jerusalem, Israel

See also Competition, Political; Electoral Behavior; Electoral Campaigns; Electoral Geography; Electoral Systems; Electoral Turnout; Participation; Parties; Party Organization

Further Readings

Dalton, R. J., & Wattenberg, M. P. (Eds.). (2000). *Parties without partisans: Political change in advanced industrial democracies*. Oxford, UK: Oxford University Press.

Galderisi, P. F., Marni, E., & Lyons, M. (Eds.). (2001). *Congressional primaries and the politics of representation*. Lanham, MD: Rowman & Littlefield.

Gallagher, M., & Marsh, M. (Eds.). (1988). *Candidate selection in comparative perspective: The secret garden of politics*. London: Sage.

Hazan, R. Y., & Pennings, P. (Eds.). (2001). Democratizing candidate selection: Causes and consequences [Special issue]. *Party Politics, 7*(3), 267–380.

Hazan, R. Y., & Rahat, G. (2006). Candidate selection: Methods and consequences. In R. S. Katz & W. J. Crotty (Eds.), *Handbook of party politics* (pp. 109–121). London: Sage.

Hazan, R. Y., & Rahat, G. (2010). *Democracy within parties: Candidate selection methods and their political consequences*. Oxford, UK: Oxford University Press.

Mayer, W. G. (2001). Primary elections. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 12011–12014). Oxford, UK: Elsevier.

Narud, H. M., Pedersen, M. N., & Valen, H. (Eds.). (2002). *Party sovereignty and citizen control: Selecting candidates for parliamentary elections in Denmark, Finland, Iceland and Norway*. Odense: University Press of Southern Denmark.

Ranney, A. (1981). Candidate selection. In D. Butler, H. R. Penniman, & A. Ranney (Eds.), *Democracy at the polls: A comparative study of competitive national elections* (pp. 75–106). Washington, DC: American Enterprise Institute.

Ware, A. (2002). *The American direct primary*. Cambridge, UK: Cambridge University Press.

Elections, Volatility

A recurring exercise in the field of electoral behavior is assessing to what extent parties' share of the vote and voter choice remain stable over time. Voters may remain loyal to the same party they voted for in the previous election or change preference by voting for another party, abstaining, or spoiling their ballots. A significant number of studies have been devoted to understanding the dynamic of such changes in contemporary democracies. The change of voters from one party to another in the election is called election volatility.

Possible changes in voters' preferences over time may be measured on the individual level or on the aggregate level of the electorate. A voter may change preferences during an election campaign, or over a longer period involving at least two elections. The traditional way of identifying the level of permanence/change in electoral preference on the individual level is quite simply asking voters. For this reason, election surveys are the fundamental source for picking up permanence/change in preference on a microlevel.

For electoral studies on the aggregate level, the source of data is the final result of the race (parties' vote, turnout, and spoiled votes) for a certain electoral unit (national or subnational) in at least two elections. Easy access to election results throughout the world may be the main reason why studies about electoral changes on the macrolevel have been conducted with such frequency over the past few years.

Many indices have been proposed to measure permanence/change in electoral preferences. The most fortunate, and one widely used at present in studies on comparative party systems, is the *volatility index*, created by Mogens Pedersen (1979, 1980). The term *volatility* first appeared in chemistry and refers to the tendency of a substance to vaporize. It has also been defined as a measure of how readily a substance vaporizes. The notion of a change in physical state captures the fundamental idea of the index, which is to quantify the intensity of change in party preferences when two elections are compared.

The Index

The volatility index (or total volatility, TV) measures the level of aggregate electoral change over two consecutive elections. Although the index is traditionally calculated for the total number of votes received by parties, it can also be calculated including total turnout and spoiled votes. Its formula is

$$\mathrm{TV} = \frac{1}{2}(|PaV| + |PbV| + |PcV|),$$

where V represents the percentage difference of votes (or seats) for each party in two consecutive elections.

The volatility index is calculated as follows: The percentage of votes (or seats) that a party received in an election is subtracted from the percentage of votes obtained by this same party in the preceding election; the difference indicates the change, and the (−) or (+) signs reveal the decline or growth of a party, respectively. The next step is to add the result of this operation (not considering the sign) and divide by two. The value of the index expresses the total number of votes lost by the parties whose votes fell from one election to the next or the total number of votes gained by the parties whose votes grew in the same period. For example, a volatility index of 10.0 indicates that the combined share of the vote of all the parties whose vote fell was 10 percentage points and, consequently, that the parties whose vote grew, taken together, increased their share of the vote by the same figure.

Theoretically the index can vary from 0 to 100, although these values could only occur in empirically improbable cases. The index is zero when the parties have identical percentages of votes (or seats) in two successive elections. The index is 100 when a radical revolution in voters' preferences takes place between two elections: The parties that ran in Election 1 receive no votes in Election 2.

One must be careful on two counts when interpreting the volatility index. The first is that even in a situation in which there are no changes in voters' preferences there may be a certain degree of volatility. The reason is simple: New voters may have a voting pattern that is different from that of traditional voters. In such a situation, volatility is derived from the difference between the votes cast by new voters and those that have been cast by voters who have since died (see Bartolini & Mair, 1990, p. 21).

The second count is regarding the relation between changes in the sphere of individual preferences and aggregate volatility. A party might maintain the same percentage of votes in two consecutive elections on aggregate but not on the microlevel. For example, imagine that two parties (A and B) have each received 50% of the votes in two consecutive elections but that all the voters of A in Election 1 voted for B in Election 2 and vice versa. In spite of the fact that all the voters changed their vote on the microlevel, total volatility was equal to zero.

Scott Mainwaring and Edurne Zoco (2007) carried out a wide-ranging comparative study on the subject of electoral volatility in 47 countries in the post-1945 period. The mean values found for each country varied significantly: The lowest volatility found was that of the United States (3.3), while the highest was for Ukraine (59.2). The mean volatility of the countries studied is 23.2; the median is 17.0. The authors investigated whether the variation in volatility is associated with a set of variables. Two of them have some statistically significant association with volatility: fragmentation of the party system and duration of the democratic regime. They argue that fragmentation is positively associated with volatility, while duration of the democratic regime is negatively associated with it. Put differently, countries with low fragmentation and a democracy of long duration would tend to have less volatile party systems (Mainwaring & Zoco, 2007, pp. 164–165).

Bloc and Intrabloc Volatility

Stefano Bartolini and Peter Mair (1990) proposed the calculation of *bloc volatility* (BV) and *intrabloc volatility* (IV). Unlike the total volatility index, which is based on gains and losses on the part of individual parties, BV and IV take as their starting point a set of parties, such as government parties × opposition parties, right × left, religious parties × secular parties, and so on. Bloc volatility has the following mathematical expression:

$$\mathrm{BV} = \frac{1}{2}\sum P(aV + bV + cV) + \sum P(dV + eV + fV),$$

where $P(aV + bV + cV)$ represents the percentage of votes gained (or lost) by the parties of this bloc in relation to the bloc $P(dV + eV + fV)$. Bloc volatility is part of the total volatility and evinces the exchange that occurs between the two blocs.

Intrabloc volatility (IV) is equal to the difference between the total volatility and the bloc volatility. Hence,

$$\mathrm{TV} = \mathrm{BV} + \mathrm{IV}.$$

Intrabloc volatility can be measured by the sum total, for each bloc, of the parties with an algebraic sign different from that of the bloc as a whole. For example, let us imagine two blocs with three parties and the following individual volatilities: (+6, −2, −1), (−5, −2, +4) and TV = 10. In this case, BV and IV would be, respectively,

$$\mathrm{BV} = \frac{|+6 - 2 - 1| + |-5 - 2 + 4|}{2} = 3$$

and

$$\mathrm{IV} = |-2| + |-1| + |+4| = 7(10 - 3).$$

An Index in Search of a Concept?

There is not much controversy about the capacity of the volatility index to detect changes in the party system on the aggregate level. The index is good both for diachronic comparisons within a country and for comparisons between countries. Hence, it can be said that a country has a higher volatility index than another or that volatility in a certain country fell, for example.

As Gary Goertz (2005) notes, a classic effort of the social sciences is producing indices that may make certain concepts operational in the best way possible. But the incorporation of the term *volatility* into political science did not follow this path. The term began to be used to designate a recently created index. In other words, the term *volatility* is more associated with a measurement than with a concept. This brand of origin becomes clear when one observes the definition of volatility, which may be found in several texts on the theme: Volatility is the net change within the electoral party system resulting from individual vote transfers.

Bartolini and Mair (1990) have pointed out the conceptual fragility of the volatility index:

> A variety of different theoretical meanings have recently been attributed to the concept and measure of electoral volatility. For example, evidence of volatility has been read as an indicator of the "integration" or "catchall" nature of party support, as a symptom of realignment and/or dealignment; as an indicator of cleavage strength and persistence; as expressing an "issue-effect" *vis-à-vis* the basic vote of parties, and so on. (pp. 23–24)

For this reason, the volatility index has been used to measure different political processes: electoral dealignment/alignment (Raul Madrid, 2005; Pedersen, 1980), the freezing of political cleavages (Bartolini & Mair, 1990), and the institutionalization of the party system (Michelle Kuenzi & Gina Lambright, 2001; Mainwaring & Zoco, 2007).

This conceptual ambiguity, added to some other difficulties inherent to any interval measurement—when used for the analysis of institutions—has left open a series of questions. What level of volatility would allow one to consider a party system to be institutionalized (or not)? How should one interpret low volatility in a party system with low institutionalization? And, conversely, how should one interpret high volatility in a consolidated party system? In substance, what does it mean to say, for instance, that the average volatility for Colombia is 12.5, while that of Canada is 11.9?

In spite of the limits pointed out, the volatility index is one of the most widely known and used indicator in political science. It has helped in identifying tendencies in longitudinal studies of a single country and contributed to the effort of comparing several party systems—a rare success story for an indicator.

Jairo Nicolau
Institute of Social and Political Studies,
Rio de Janeiro State University
Rio de Janeiro, Brazil

See also Election Research; Electoral Behavior; Electoral Campaigns; Electoral Systems; Electoral Turnout

Further Readings

Bartolini, S., & Mair, P. (1990). *Identity, competition and electoral availability: The stabilisation of European electorates 1885–1985.* Cambridge, UK: Cambridge University Press.

Kuenzi, M., & Lambright, G. (2001). Party system institutionalization in 30 African countries. *Party Politics, 7*(4), 437–468. Retrieved December 6, 2010, from http://ppq.sagepub.com/cgi/content/abstract/7/4/437

Lane, J.-E., & Ersson, S. (1997). Parties and voters: What creates the ties? *Scandinavian Political Studies, 20*(2), 179–196.

Madrid, R. (2005). Ethnic cleavages and electoral volatility in Latin America. *Comparative Politics, 38*(1), 1–20.

Mainwaring, S., & Zoco, E. (2007). Political sequences and the stabilization of interparty competition: Electoral volatility in old and new democracies. *Party Politics, 13*(2), 155–178. Retrieved December 6, 2010, from http://ppq.sagepub.com/cgi/content/abstract/13/2/155

Pedersen, M. (1979). The dynamics of European party systems: Changing patterns of electoral volatility. *European Journal of Political Research, 7*(1), 1–27.

Electoral Behavior

In representative democracies, voting is one of the fundamental acts that allow citizens to express their belonging to a political community and to make decisions about their own future. Analyzing electoral behavior is one of the oldest and most productive domains of research in political science. Research on electoral behavior contributes to our understanding of how democratic systems function and develop over time. It does this by analyzing the answers to the following questions: Whom do people vote for? Why do they do so? What do they want to achieve by doing so?

This research has developed over many years using three main approaches: sociological, psychosociological, and rational choice. It should be remembered nonetheless, in France particularly, that the first studies concentrated on electoral geography and were based on analysis of aggregate electoral data at different geographical levels and that this subdiscipline continues to be of great interest.

The sociological approach endeavors to explain the electoral behavior of individuals according to their position in society and their belonging to particular social groups. It studies the social determinants of voting behavior. The psychosociological approach studies the individual attitudes of voters on which they base their electoral choices. The rational choice approach developed from economic theories of rationality. Voters decide which way to vote using a cost–benefit calculation, depending on their preferences and what the different parties and candidates have to offer. These approaches are discussed in detail as follows.

The Sociological Approach Based on Social Cleavages

Sociological research on voting behavior has been developed since the 1950s as a result of the increasing presence of major opinion poll surveys based on large samples of voters. These surveys allowed a great deal of individual data to be gathered and submitted to complex quantitative processing. In many countries, strong statistical relationships appeared between the social characteristics of individuals and the way they voted. Among the main social variables used and whose effect has been measured, class and religious belonging were initially the most frequently analyzed, and they also appeared to be the factors that had the most determining effect.

Seymour Lipset and Stein Rokkan's theory of cleavages looked at the history of European societies to identify the main stable social cleavages at the roots of political and partisan cleavages. They exerted a huge influence on research into voting behavior by showing the profound and lasting nature of the major social determinants of voting. Their research has provided solid results on this question. The study of class voting was thus well developed, as was the study of the relationship between religious belonging and voting behavior. Thus, for example, while class constituted the most important electoral cleavage in Great Britain, religion seemed to be the variable with the greatest impact in both Germany and France, though class also played a role but to a lesser extent.

More recently, this approach has been criticized in several ways. It was observed that the ability of these two variables to explain voting behavior was becoming weaker in accordance with a long-term trend. Thus, the permanence of class voting, particularly in Great Britain, but also more broadly speaking in Europe, became debatable as a result of analyses that showed an underlying decrease of the left-wing working-class vote and the right-wing middle-class vote. Many explanations for the decline of class voting have been suggested. Some of these point to changes in social structures in Western societies and to both the decrease and fragmentation of the working class. Others stress the individualization of these societies and the increasing autonomy of individuals in relation to the groups they belong to. This hypothesis is also suggested for the religious factor, as European societies have become increasingly secular.

Others again stress changes in the political parties themselves rather than changes among the electorate. According to this hypothesis, it is no longer in the interest of the parties to present themselves as "class-based parties" for the purpose of winning an election. This has contributed to weakening the coherence of the different classes as self-conscious collective actors. As a response to these criticisms, other researchers have relativized or denied the "end of class voting." According to them, current surveys show that the decrease of class voting does not imply its disappearance. The operationalizations designed to measure it have also been criticized. It is argued that the use of more detailed and more complex social and occupational classifications would reveal that the relationship between belonging to a social group—and not necessarily to one of the two major classes conceptualized by Marxism—and voting behavior remain significant. Finally, the ideological supposition according to which the working-class vote, for example, can only be classified as a class vote if it is a left-wing vote can be called into question. For example, in France in 2002, working-class voters voted more than any other social group for the extreme Right. How can such a vote be qualified according to the class voting theory?

The same is true for the religious variable. Although surveys (notably in Europe) have revealed a decrease in both religious attachment and the relationship between religious belonging and voting behavior, this relationship is far from having disappeared. Furthermore, the increasing importance of Muslims in Europe has made this variable relevant once again in the analysis of voting behavior. In France, for example, Muslims are far more likely to vote left than right. In the same way, in the United States, presidential elections in this century still show the permanence of the strong relationship between religion and how people vote.

Finally, even if the hypothesis according to which the vote has become individualized should be taken seriously, it nonetheless merits a certain degree of relativity. In particular, although the role played by certain social variables in voting behavior is diminishing, the role played by other social variables is emerging, reappearing, or indeed gaining importance. This applies, for example, for gender and age. Furthermore, in a number of European countries, the regional vote, which is a vote based on a sense of identity, has been strongly expressed for several years, such as in Scotland, in Catalonia, in the Basque country, and in Flanders. In the same way, the ethnic or racial vote is not receding either and may, indeed, be increasing. In the 2008 American presidential elections, 95% of African Americans voted for Barack Obama as compared to 43% of Americans of European descent.

Whatever the tendency toward the individualization of Western societies, voters remain individuals belonging to different social groups and networks. Individuals have multiple identities, and one or another of these may more or less dominate and be mobilized depending on the period, the events, and what the parties and candidates are offering. Although the sociological approach has real limits—and in particular it has not always sufficiently concerned itself with what politics has to offer, at times reifying a given determinant and overestimating its permanence or its centrality—it remains nonetheless an essential approach in the analysis of electoral behavior.

The Sociopsychological Approach

In the 1950s, the publication of *The American Voter* brought about a turning point in the study of electoral behavior. The starting point for the model used—known as the Michigan model—was, on the one hand, the apparent contradiction between the absence of political sophistication among the electorate and the instability of their opinions on different issues and, on the other hand, the fact that the vast majority of voters constantly voted for the same party throughout their lifetime. This model, which was founded on the importance of political socialization (familial, professional, ethnic, regional, etc.) and on the way in which political choices and preferences develop among individuals, has produced the concept of party identification to explain electoral behavior. Partisan identification is of an affective nature. It is politically stable, and this stability may even become stronger with age. Individuals often reproduce the same partisan preferences of their parents. This partisanship creates a basis for political identity and provides cues for evaluating political events, candidates, issues, and subsequently for voting preferences.

This model also has a strong heuristic value. However, it has nonetheless been criticized for many reasons. It has been judged to be tautological and nonexplicative—voters who vote for a given party say they feel close to the party as a result of voting for it and not the other way around. Furthermore, more recent studies have shown a strong decrease in partisan identification in several countries. Various theses on the modernization and individualization of our societies have highlighted the fact that it is less useful for a more and more sophisticated electorate to use partisan identification as a shortcut in deciding how to vote. Other studies have measured increasing levels of individual electoral volatility and have highlighted the fact that the partisan de-alignment phenomenon is part of a general process of political change that is transforming the relationship between voters and parties. Finally, it has been noted that this concept has always been more relevant in two-party systems than in countries where the system is multipartisan and unstable.

The psychological and, in particular, the affective dimension of the concept should be put into perspective although it does still exist for a certain category of voters. However, a little vaguer but perhaps more heuristic notion of partisan proximity, but also possibly of partisan distance, remains useful. For many voters, it represents an important way to locate themselves in the political landscape. The general dynamic of political modernization tends to individualize electoral behavior. However, this does not mean that all citizens have the same relationship to politics, the same level of sophistication, and the same interest in or information about politics. Some people have a greater need of shortcuts and symbolic clues to help them decide how to vote. In continental Europe, where party systems are frequently multipartisan, the notions of left and right are often more useful tools to study the orientation and stability of political preferences than the notion of party identification, as used in Great Britain or the United States. The notions of Left and Right can play a comparable role. Studies have shown, furthermore, that the importance of the voter's relationship to the party—or to the Right or the Left—is not eliminated by the increasing importance of the leading candidates' personalities. This is all the more true as the major candidates are designated and supported by political parties.

Attitudes of an affective nature in the relationship between the voter and the political system do not, however, only concern the relationship to political parties. They also concern the candidates and other political personalities. The dynamic at work in most democracies shows the increasing importance of the personalization of politics, in particular in presidential or semipresidential systems. The role played by the media, by opinion polls, and by expert opinion is in strong competition with the role played by the parties and favors the personalization of politics. The choice of a leader includes an affective dimension to the extent that, at a given moment, the leader incarnates responsibility for the destiny of the political community. Studies have shown that the governmental credibility of a leader plays an important role in decisions on how to vote. This is particularly true at times of serious national crisis, whether it be economic, international, or political. The notion of the leader's credibility, which is an important dimension for electoral choice, cannot be separated from the notion of confidence, which is partly of an affective nature. This is why, when designating candidates, the major parties must take the charismatic aspect of their personalities into account. Admittedly, studies have undermined the hypothesis according to which voters essentially make a decision based on the personal image of a candidate and independently of their personal political orientations and of those of the candidate. Conversely, however, the candidate's personal image cannot be considered to be secondary. Political leaders have become electorally important in their own right by personifying the policy platforms of their respective parties. This personification necessarily contains a psychological dimension even if this is not the only dimension.

The psychosociological approach to voting behavior does not merely concern the study of the effects of partisan identification and personalization. It also takes the voter's ideological orientations and value systems into account. Even though these systems may be complex and modifiable, they are nonetheless consistent and long lasting. In the political landscape, some of the major value cleavages reflect cleavages in the politico-ideological field. From this point of view, the role of the left–right dimension is also important. In many countries, studies have shown that two main subdimensions of attitudes provide structure to the left–right dimension. These are the economic dimension, on the one hand (economic liberalism vs. state interventionism), and the societal dimension, on the other hand (cultural liberalism or libertarianism, or in a certain sense, postmaterialism), versus cultural conservatism. A voter who subscribes to both economic liberalism and cultural conservatism is more likely to vote for the Right, whereas hostility to economic liberalism and support for cultural liberalism is more likely to foster a vote for the Left. To take this further, France is a country where there is great diversity in what is on offer politically. It has been noted in recent elections that the existence of a supply of centrist and extreme-right parties has favored the political expression of economic antiliberals and cultural antiliberals who voted for the extreme Right and economic liberals and cultural liberals who voted for the centrist parties. It seems, therefore, that the left–right dimension is in reality not a continuum but a complex structure made up of different ideological dimensions that are expressed in different ways depending on the period and on the country. Thus, in many European countries, the increasing importance attributed by the socialist left to the values of cultural liberalism and antiracism, which is widely shared by the well-educated middle classes, encouraged a certain number of working-class voters to change their electoral behavior. These were voters who, up until then, had been attached to the left because of their economic antiliberalism positions but whose xenophobic tendencies were increasing. They availed themselves of the presence of an extreme-right party that combined the two types of antiliberalism in its ideology and voted for it.

Similar to the sociological approach, the psychological and sociopsychological approaches have kept heuristic power. Furthermore, it should be noted, as this last example shows, that the two approaches are complementary to the extent that there are strong relationships between social groups and value systems.

The Rational Choice Approach and Short-Term Elements

The rational choice approach has more recently influenced the study of voting behavior. Its point of departure is the notion of a cost–benefit calculation. Voters vote in accordance with their preferences for the party that will best satisfy them. Based on a microeconomic decision-making model, illustrated notably by Anthony Downs, this approach differs from the sociological and psychological approach in that it considers the voter to be a rational individual. According to this model, citizens vote for the party they believe will provide them with more benefits than the other parties. Parties offer policies; voters look at these policies and decide which of them will provide maximum benefits for them and vote accordingly. Parties and candidates, on the one hand, and voters, on the other, are in a constant state of interaction. Parties constantly adapt to the demands of the voters in an attempt to conquer or maintain power, which, according to this theory, is their essential and indeed only aim.

This approach has been very strongly criticized in different ways. It has been argued that the notion of interest or utility is not directly applicable to the relationship of citizens to politics, as their expectations are expressed in a complex way both in terms of material and symbolic retributions and also in terms of precise demands and attachment to values. It becomes necessary to attribute such a broad meaning to the notion of preference that the use of the notion of utility becomes problematic. Certain critics have gone so far as to claim that this approach has no real value. The result has been a veritable battle between two schools of thought.

And yet this model holds many advantages. It highlights the importance of the process of interaction between voters and parties. This approach stresses that any study of voting behavior cannot take only the voter into account without also including the political system, the party system, and the set of alternative propositions (set of policies) proposed by the parties and candidates. Here, the role played by what the parties have to offer is essential. The parties frame what is on offer politically, and the choice made by the voter is largely constrained by the issues that the parties themselves have chosen to highlight.

Another advantage of this model is that it places the voter/individual at the heart of the process and considers that, whatever their position in society, voters are capable of working out their electoral choices themselves. What is involved is an individual decision-making process. A significant contribution offered by this approach is that it makes a real attempt to build complex statistical models to enhance understanding of electoral behavior. The development of this approach has contributed in a certain sense to the current tendency in electoral analysis to show a general decline of the long-term determinants, a growth in the importance of short-term attitudes, and a fragmentation of voter choice (diversity of decision-making processes). This tendency notably benefits research on issue voting and economic voting.

Analyses of issue voting partly stem from the rational choice approach to the extent that they deal with the interactive political relation between supply and demand and presuppose that the voter is relatively sophisticated. Two types of issues are usually identified: position issues, where the parties give priority to different issues, each in relation to the other, with these issues thus being in competition with each other, and performance issues (valence issues), where the parties pursue objectives aspired to by all (e.g., a reduction in the rate of unemployment) and try to demonstrate their competence in this respect as a means to reach them.

In terms of issue voting, the rational choice approach has produced spatial theories on voting behavior. The proximity theories of voting place voters and parties in a political landscape. According to these theories, voters feel close to the parties they vote for and distant from those they don't vote for, and there is a measurable distance between the voter's position on policies and issues and the party's equivalent positions. There is a distance between the voter and the parties on offer, the function of this distance being to identify which party a voter will choose in a given election. The "proximity" function, which has long been accepted for modeling party and self-placement data, has been challenged by directional theories. Based on an analysis of the notion of proximity, the rational choice models concluded that it was in the interests of the parties to adopt the position of the median voter. Supporters of the directional thesis bring not only the choice of an issue by the voter into play but also the (psychological) intensity of the preference for the issue in question. Here, the rational choice model includes recourse to the psychosociological one. Indeed, within this model, voters may prefer a party that has a very strong position on the issue in question, even if it is not the party the voter feels closest to. By the same token, it may be in the party's interest to adopt a strong position on the issue and not necessarily to attempt to position itself in relation to the median voter. Under these conditions, parties do not necessarily converge toward the center of the political landscape. Today, this family of models represents the formalized attempt to draw up a synthesis between the different approaches encompassing both the rational choice elements of the proximity model and the psychological processes of the directional equivalent.

All analyses of issue voting pose problems of a more general nature. First, their specific contribution cannot be evaluated without, at the same time, taking the contributions made by other types of variables that affect voting behavior into account. For example, a voter who feels very close to a given party will tend to attribute the best mark to this party on all the issues proposed. Thus, in this model, the question of causality (what comes first?) has to be addressed. Second and more important, preferences on issues are not independent of the more clearly articulated and long-lasting value systems of voters or of their social position and socialization.

Therefore, both the short-term (new issues) and the long-term (stable and consistent ideological alignments) factors must be taken into account. From this point of view, the analysis of issue voting cannot be carried out independent of the analysis of values and attitudes. However, a new issue can, for example, disturb or contribute to the reconfiguration of different dimensions of a voter's value system or indeed reverse the priority given by the voter to one dimension over another within the system. It can update the voter's long-term predispositions (concept of funnel of causality).

Short-term analyses also cover economic voting. They are based on the hypothesis according to which an analysis by the voter of the economic performance of those in power or simply the voter's perception of the economic situation and its development has an influence on his or her vote. Analysis of economic voting poses several problems even if certain results remain convincing. First, the results obtained are not always consistent, depending on the country and period of time involved. Second, retrospective judgment analyses on policies implemented, particularly in the economic field, show that, similar to judgments on issue voting, they are not independent of voters' partisan proximities and their political orientations and therefore of the political orientation of the incumbent government.

Conclusion

The study of voting behavior is an extremely complex affair for all the reasons described above but also because political systems evolve over time and according to country, and voters adapt to the changes in these systems. Whatever the attempts to universalize the different theories, the significance of the vote differs from one society to another, from one historical phase to another, and from one voter to another. The studies available on societies other than North American and European ones (which are still too scarce) show that the major approaches used here do not always work very well in other systems. Thus, in democratic societies in the making, with embryonic party systems and without any real unity or national awareness, and in societies that are still traditional, the social determinants that produce the vote may be quite different. Many factors can strongly influence the vote, such as belonging to an ethnic, religious, or village community; belonging to a client-type network; the existence of mafia-type systems; and so on. Research must therefore extend itself in the future to include the whole range of electoral behaviors.

Finally, research on institutional aspects and the rules that organize voting must also be developed. This is notably the case for the effects on electoral behavior of voting systems, the size of constituencies, the political and institutional issues at stake during a given election, and the material organization itself of the vote.

Gérard Grunberg
CNRS/Sciences Po Paris
Paris, France

See also Class, Social; Cleavages, Social and Political; Electoral Geography; Electoral Systems

Further Readings

Campbell, A., Converse, P. E., Miller, W. E., & Stoke, D. E. (1960). *The American voter*. New York: Wiley.

Dalton, R. (2008). *Citizen politics: Public opinion and political parties in advanced industrial democracies* (5th ed.). Washington, DC: CQ Press.

Dalton, R., & Klingemann, H. D. (Eds.). (2007). *The Oxford handbook of political behavior*. Oxford, UK: Oxford University Press.

Downs, A. (1957). *An economic theory of democracy*. New York: Harper.

Evans, J. (2004). *Voters and voting: An introduction*. London: Sage.

Lewis-Beck, M., Jacoby, W. G., Norpoth, H., & Weisberg, H. (2008). *The American voter revisited*. Ann Arbor: University of Michigan Press.

Shanks, M., & Grofman, B. (1999). *A unified theory of voting*. Cambridge, UK: Cambridge University Press.

Thomassen, J. (2005). *The European voter: A comparative study of modern societies* (ECPR series). Oxford, UK: Oxford University Press.

Electoral Campaigns

Electoral campaigns are the means by which political parties, alliances, coalitions, and candidates convey their policies and election programs to the electorate in an election contest. Electoral campaigns are essential in a democratic society that holds national and local elections on a periodic basis. International standards governing the environment, conduct, and principles involved in electoral campaigns in democratic elections have evolved over the past several decades.

In order to create an environment conducive to the organization and conduct of free and fair elections, a number of checks and balances have been developed over time to assist parties and candidates in ensuring a level playing field during the campaign period. Similarly, measures have been developed to avoid or reduce the threat of intimidation or violence against candidates, parties, or their supporters. In addition to provisions dealing with election campaigns in the given legal framework, a code of conduct aimed at parties, candidates, and their supporters is often recommended. In a similar vein, a code of conduct to guide the media, particularly the publicly owned media, is always recommended.

Nature of Electoral Campaigns

Many electoral legislative schemes stipulate a fixed period during which electoral campaigning should take place. The stipulated period may range from 14 days to 11 weeks or even longer. The campaign period may be required to end 24 or 48 hours, or even some days, before polling day. The electoral law or the code of conduct for parties, alliances/coalitions, and candidates sometimes impose penalties for campaigning outside the official designated period.

Electoral campaigns flourish best where there are constitutional guarantees of freedom of association and freedom of speech. These fundamental rights are essential to the electoral campaign environment, as was vividly shown to the world in Zimbabwe's run-off election in June 2008 when the opposition candidate, Morgan Tsvangirai, had to withdraw from the contest because of disruption to his campaign rallies, intimidation, and violence against his supporters.

Methods of Electoral Campaigns

There are a number of electoral campaign methods, including rallies, motorcades, radio broadcasts, television broadcasts, debates, party manifestos, advertisements, house-to-house canvassing, and the use of banners, fliers, buntings, billboards, shopping bags, clothing, and theater plays. In many countries, the issuance of permits to hold rallies and motorcades resides with the police authorities, who can act in favor of the ruling party. The issuance of permits to hold rallies or motorcades often operates more smoothly when it is done by the electoral management body in cooperation with the police authorities as in Gambia and Sierra Leone.

Media access, particularly the publicly owned media—electronic and print—is sometimes regulated by specific legislation and/or by a media code designed to regulate election campaign reporting. The publicly owned electronic media may allow free airtime in accordance with an agreed formula (based on equal or equitable or fair allocation free time). In many new and emerging democracies, the public media find it difficult to follow the rules relating to the allocation of free time to the opposition. The private media sometimes enter into voluntary conduct aimed at developing balanced reporting during an electoral campaign, but this approach is not widespread.

The behavior of the media can have a significant impact on leveling the playing field for the contestants in an electoral campaign. Fair access is only one dimension of the media environment. There is also the issue of whether or not the media reports campaign news in a balanced and nonpartisan way and whether misstatements are corrected with equal prominence as the original report.

Manifestos

It is customary for political parties, alliances, coalitions, and candidates, including independent candidates, to publish an election manifesto or platform during the electoral campaign. Such an election manifesto or platform sets out the program that the parties or candidates intend to put across to the voters as the one they would implement if they were elected. This proposition suggests that manifestos in national elections are about issues affecting voters and the country. However, often, perhaps too often, the campaigns are run on the personality of individual candidates rather than the issues raised by the manifestos.

The extent to which a party's or candidate's manifesto or platform is binding may be political rather than legal. Parties or candidates who do not honor their manifestos' pledges may pay a political price at the next election. Of course, these programs may be modified to meet changed circumstances, but such changes should be properly explained to the electorate. There have been cases where elections have been held without political parties, as in Uganda until recently and Swaziland, and there are a number of countries where independent candidates are not allowed to contest elections, as in Tanzania.

Incumbency

The term *incumbency* refers to the party in power, that is, a party or the parties forming a government at some level or to an individual holding an office. An incumbent authority usually attracts advantages in a campaign by virtue of the office held, unless steps are taken to reduce the incidence of abuse of such an office. To avoid incumbency abuses, some countries, most notably India, have introduced strict measures by way of a code of conduct governing incumbent parties and candidates during electoral campaigns. These measures, for example, restrict the combination of ministerial functions with party political campaigning and forbid the use of public resources for campaign purposes. Bangladesh's attempt to reduce the influence of incumbency on general elections was to require the government to leave office at the end of its term and turn over the role of government to a caretaker government to oversee the preparation and conduct of the elections. Regrettably, that approach has not worked well and not only has it led to the disintegration of the democratic path to holding elections but also in part has brought down the entire system of democracy and brought in an army-backed regime.

Level Playing Field

The phrase *level playing field* has taken hold with respect to electoral campaigns and is used to describe the treatment meted out to competing contestants during an electoral campaign. Often, complaints are triggered by stakeholders when they consider that the election environment does not give rise to a level playing field. This situation arises frequently with access to the media and to the use of public resources. The negative aspect of a level playing field sometimes comes into play when the police authorities delay or refuse to grant permits for election rallies.

Campaign Finance

Dealing with campaign financing may range from no regulation by the state or the election management body (EMB) concerned to tight controls either by the state or by the EMB. Such controls may include limits on maximum expenditure by parties and/or candidates, party-sponsored or independent maximum public contribution, or maximum contribution by individuals or companies. There may also be strict rules governing disclosure of the amount of private contributions received during the electoral campaign. Such disclosure is often open to public scrutiny, and the contributions are subject to audit by reputable auditors. Campaign finance reporting should be required at a stipulated period before and/or after polling day. Failure to meet the stipulated deadline usually incurs significant sanctions, and any material breach of the campaign finance rules may result in forfeiture of the elected office by a successful candidate.

Role of the Media

The media play the most important role in conveying to the electorate the message of parties and candidates. The electronic media has been playing an increasingly prominent role through television, radio, and the Internet (mainly in the developed world). In many new and emerging democracies, the radio has the greatest reach with respect to comprehensive countrywide coverage (although in a few countries even the use of radio batteries during a campaign period have proven to be beyond the means of rural dwellers).

Often, a distinction is made for electoral campaign purposes between publicly and privately owned electronic media. This is because the ruling party frequently influences those public bodies to operate in a partisan manner. This behavior is often mitigated by codes of conduct for the media whereby a schedule for free time broadcast by parties and candidates is regulated by statute or voluntarily. Free time allocation is done by a formula based on fair, equal, or equitable distribution. The basic rule of good practice is that, in addition to fair access to the electronic media, the reporting of election news and the coverage thereof should be balanced.

The print media also play an important role in many campaigns. A similar distinction between public and privately owned print media and code of conduct may be useful in creating a level playing field with respect to access between the ruling party and others. However, in addition to fair access to publicly owned print media, the privately owned media should have access to print paper and to distribution outlets as in normal times outside campaign periods.

Security

The security of party functionaries and candidates is pivotal during the period of election campaigning. Some electoral legislative schemes, supported by codes of conduct for parties, their supporters, and candidates, provide the security framework for election campaigns. Occasionally, security measures that were designed for election campaigns do not follow through to the immediate postpolling period as happened during the post-election violence in Kenya in 2007–2008. Equally disturbing, the election campaign security may break down through the partisan behavior of the security forces either at the instigation of the governing party or with their complicity, as in the run-off presidential election in Zimbabwe in June 2008 between Robert Mugabe and Morgan Tsvangirai. Inadequate campaign security in the past, and even recently, as, for example, in December 2007, when Benazir Bhutto was assassinated at a campaign rally in Pakistan, can profoundly and adversely affect election campaigns and may even influence the results. The security forces employed during election campaigns should be properly trained and ensure adequate security to all contestants without partisanship during an election campaign.

Intimidation and Violence

When perpetrated during an election campaign and designed to interfere in any way with an opponent's legitimate campaign activities, intimidation and violence should be condemned and sanctioned strongly under the election law and/or the code of conduct for parties, their supporters, and candidates. Credible reports by independent election observers, international and domestic, and other stakeholders described a campaign environment in Zimbabwe for the June 2008 run-off presidential elections of relentless and continuing violence and intimidation that resulted in many deaths and injuries. According to many stakeholders, the sustained vicious intimidation and violence directed mainly at supporters of the opposition could be attributed to premeditated planning by the ruling-party to prevent the opposition contesting the run-off elections. Rallies planned by the opposition were either denied the required police permit or, when the required permit was granted, disrupted by ruling-party activities. The consequence of the unlawful acts against the opposition was the withdrawal of their presidential candidate, leaving the incumbent presidential candidate unchallenged at the polls.

Conclusion

In recent years, the spotlight in election organization has been very much on campaigns. This has led to more political parties adhering to codes of conduct and subscribing to greater transparency in campaign financing. The growing acceptance of domestic and international observation at national elections campaigns is enhancing the transparency in election planning and conduct. Notwithstanding the notable improvements being made, certain undesirable practices linger on in emerging democracies in the African Union countries and in countries in Eastern Europe (including Russia) and in central Asia. These undesirable campaign features include undue influence on EMBs, abuse of incumbent advantages, not having a level playing field in respect of access to publicly owned media, and use of public resources by ruling parties.

Carl W. Dundas
Dundas and Associates Ltd.
Furzton, Milton Keynes, United Kingdom

See also Election Observation; Election Research; Electoral Geography; Electoral Systems; Electoral Turnout

Further Readings

Dundas, C. W. (1996–1998). *Compendium of election laws, practices and cases of selected commonwealth countries* (Vol. 1, Parts 1 & 2, Vol. 2, Parts 1 & 2). London: Commonwealth Secretariat.

Dundas, C. W. (1999). *Election management bodies: Constitutive instruments.* London: Commonwealth Secretariat.

Dundas, C. W. (2006). *Improving the organisation of elections: A 2006 perspective.* Kingston, Jamaica: Ian Randle.

Reynolds, A., Reilly, B., & Ellis, A. (2005). *Electoral system design: The new international IDEA handbook information unit.* Stockholm, Sweden: International IDEA.

Wall, A., Ellis, A., Ayoub, A., Dundas, C., Rukambe, J., & Staino, S. (2006). *Electoral management design: The international IDEA handbook.* Stockholm, Sweden: International IDEA.

Electoral Geography

The electoral geography hypothesis is that electoral behavior is related to place—neighborhood, community, town or city, region, or country—and cannot entirely be explained by considering individual voters, and nonvoters, as possessing characteristics on an atomized and isolated basis. It claims that geographical and contextual effects are significant and that the process of day-to-day personal communication and social interaction in the community—or in a variety of communities—influences voting patterns. As explained by John Agnew (2002), localities are the arena for political socialization and produce the settings for much interpersonal social interaction. The general tensions within society are actualized within localities, influencing the relative weight and various meanings attached to social divisions and hence the appeal of different ideologies—which are not the same everywhere. Political movements then use geography in their search for electoral and other support, using claims relating to localities, regions, and so forth to mobilize that support.

Geographical Effects on Electoral Behavior

There is a long history of debate about whether such geographical effects on electoral behavior do in fact exist. Opponents of the electoral geography hypothesis have contended that once all the variables affecting individual electoral behavior have been fully identified and properly specified, there is no role for geographical and contextual variables to have any independent effect. Furthermore, research on the existence and nature of neighborhood effects depends on the existence of usable neighborhood data: As a result, it is not necessarily easy to investigate the various levels at which they may exist.

Neighborhood effects in voting patterns were first demonstrated in 1937 in Sweden by Herbert Tingsten, who showed that the more working class the area, the more likely the working-class voter was to cast a socialist vote. Much subsequent work in this area has focused on the United States and the United Kingdom, perhaps because the political effects of the links between electoral geography and political representation are more immediately evident under majoritarian electoral systems. Support for the electoral geography hypothesis has been more extensively demonstrated in the work in the UK, over many years, of Ron Johnston and Charles Pattie, demonstrating that, in the words of William Miller (1977), "people who talk together vote together." These effects are not, however, uniform. They are, for example, more powerful in "working-class" neighborhoods than in "middle-class" neighborhoods, probably because the social circles of salaried workers are likely to be wider and more geographically varied than those of residents of poorer neighborhoods.

The impact of neighborhood effects on patterns of political representation is affected not only by the formation and evolution of social communities but also by the geographical definition of electoral communities achieved by the delimitation of electoral boundaries. The legislation, which defines how delimitation and redistribution takes place; the independence or otherwise of the authority that undertakes it; and the attempts of political parties and movements to influence it in their favor are all relevant factors. It is not possible for redistribution to produce outcomes without a bias of some form. The inherence of geographic factors in redistribution is shown in the analysis of Bernard Grofman and others (1997), breaking down bias into three component parts: (1) bias resulting from malapportionment, (2) bias resulting from different levels of turnout, and (3) bias that results from the geographic distribution of party vote shares. Although requirements on a delimitation authority to control or eliminate malapportionment through a stipulation that the electorates of all seats be equal within a stated small tolerance level are now common, statutory exceptions exist for sparsely populated areas, for example, Norway and Western Australia. In South Australia, boundary revisions are also required to take into account the geography of party or grouping vote shares and ensure that the party or grouping with more than half the two-party preferred vote sees this reflected by an elected majority in the state legislative council.

The effect of demographic changes and boundary delimitation can be accentuated by party campaigning activity, as UK Liberals and later Liberal Democrats demonstrated through successes generated by intensive campaigning in individual localities in local elections from the 1970s onward. On a larger scale, the 1997 and subsequent general election victories of the UK Labour Party also owed much to the party's efforts to "create its own geography," in Johnston and Pattie's words, through targeted geographical focus of its efforts linked to an overall increase in the effectiveness of the party machine. Electoral geography is not only determined by changing demographics and community formation and by the outcomes of boundary delimitation exercises but also by the response of political parties in targeting their resources, the intensity of campaigning, and promoting or implicitly accepting tactical voting. Furthermore, effective party management recognizes that changes in electoral geography can affect the motivation and activity of party organization and responds accordingly.

The Salience of the Electoral System

At a broader level, the overall electoral geography of a country and hence the dynamic not only of its electoral politics but of its political system as a whole are linked to the form of electoral system chosen. Cleavages can be accentuated or weakened by electoral system choice. Majoritarian systems usually have a strong inbuilt tendency to produce a seat bonus for the winning party, but there are occasions in practice where this does not happen. In Malawi in 1999, the three leading parties, United Democratic Front (UDF), Malawi Congress Party (MCP), and Alliance for Democracy (AforD), received 47%, 34%, and 11% of the national vote, respectively, which converted into 48%, 34%, and 15% of the seats. However, 82% of the UDF seats were in the southern region, 82% of the MCP seats were in the central region, and 97% of the AforD seats were in the northern region—which inevitably produced a dynamic of regional cleavage in the parliament.

By contrast, Indonesia has used list proportional representation (with differences in detail) in the democratic elections of the reform era from 1999 onward. In a political climate where the maintenance of the unity of the state was both a domestic political and a societal imperative and a priority of the international community, this type of electoral system ensured that each of the major parties secured representation from across Indonesia. Probably as a consequence, political divisions since 1999 have hardly demonstrated a regional dimension. However, mapping of the leading party in each local authority area shows a clustering of party support in different parts of the country, which would almost certainly have led to party groups in the legislature with a strong regional origin and identity. Had this happened, the likely consequence would have been the emergence of regionalism as a much stronger cleavage at the national level, with probable significant negative effects on national unity and the process of democratic consolidation. It is not only the proportion of seats gained by each party that influences the way a political system works in practice, for the geographical distribution of the leadership and elected members can be just as significant.

Future Research

While there no longer appears to be significant contest about the contribution of geographic effects to explain electoral behavior, there is still much scope for research into how they work. The study of electoral geography brings together political scientists and geographers, and there may be rich fields for future exploration both in the comparative study of electoral geographies and in the change of electoral geographies over time. Looking forward, it is a fair hypothesis that continuing changes in the way people form communities of different kinds will affect the working of the electoral geography hypothesis in practice. The spread of access to telecommunications and personal transport in many parts of the world has clearly affected the context of the evolution of communities. The growth of Internet-based social networking and the consequent advent of virtual communities may be expected also to generate new impacts on the way in which personal interaction leads to electoral decision.

Andrew Ellis
International Institute for Democracy and Electoral Assistance (IDEA)
Stockholm, Sweden

See also Data, Spatial; Electoral Behavior; Electoral Systems; Redistribution; Social Capital

Further Readings

Agnew, J. A. (2002). *Place and politics in modern Italy*. Chicago: University of Chicago Press.

Grofman, B., Koetzle, W., & Brunell, T. (1997). An integrated perspective on the three main sources of partisan bias. *Electoral Studies, 16*(4), 457–470.

Johnston, R. J., & Pattie, C. (2006). *Putting voters in their place*. Oxford, UK: Oxford University Press.

Miller, W. (1977). *Electoral dynamics in Britain since 1918*. London: Macmillan.

Tingsten, H. (1937). *Political behaviour: Studies in election statistics*. London: PS King.

Electoral Systems

One may view electoral rules for a given office as having six basic components: (1) determination of who is eligible to be on the ballot (e.g., parties only or also individual candidates), (2) internal party rules for determining who are to be a given party's candidates and/or for specifying candidate rankings within a party list, (3) specification of ballot type, (4) specification of constituencies (districts), (5) determination of election timing, and (6) rules for ballot aggregation (tallying rules). In addition, what is not to be forgotten is the seventh component of all elections—the voter—whose preferences and beliefs about how voting rules will translate those preferences into outcomes as well as the ways in which voter preferences are geographically distributed will critically affect how electoral rules operate in the real world.

Sometimes the term *electoral system* is used more broadly to include other aspects of elections and their regulation, such as rules for voter suffrage, campaign finance, campaign advertising, location of and times of access to polling stations, and so on. While all these elements are interrelated and part of the study of rules for conducting an election, in this entry, because of space constraints, the focus is primarily on ballot aggregation mechanisms. The entry examines the nature of the rules that determine electoral outcomes, identifies the most important rules used worldwide, considers alternative approaches to the study of electoral systems, and discusses domains of electoral system impact.

Typologies of Electoral Rules

Elections and the political parties that compete in them are at the heart of modern democracy. The study of elections and of the voting rules that are their engines has attracted the attention of many scholars and led to a literature, which, in the 4 decades since the publication of Douglas Rae's *Political Consequences of Electoral Laws* in 1967 (second edition, 1971), has become well developed, both empirically and theoretically. The study of electoral rules and their effects is made easier by the existence of variables that can be precisely defined and usually made quantifiable, for example, votes and seats. Because of the close connections between the study of electoral rules and the study of topics such as party systems, representation, and constitutional design, success in understanding electoral rules and their effects can have what Rein Taagepera has called a "Rosetta stone" linkage function for more general theory building in political science.

But there are also very practical reasons for trying to understand electoral system effects. Of the fundamental aspects of constitutional design, for example, along continua such as unitary versus federal systems or parliamentary versus presidential systems, electoral rules are the easiest to change. In most countries, electoral rules are not constitutionally embedded, or at least many key details are left for legislative determination. There are myriad different ways of conducting elections, and there is widespread belief, backed by empirical evidence, that choice of electoral rules matters a great deal for who gets what, when, and how—the questions that lie at the heart of all politics. Thus, manipulation of electoral rules seems an obvious tool of institutional engineering.

Candidate Eligibility/Designation of Candidates

In most polities, complex rules govern which parties and which candidate names can appear on the official ballot; for example, there may be signature requirements for nominating petitions, which require a certain number of names or a certain number of names from a certain number of different geographic regions, but some rules may be waived for parties that have demonstrated a substantial level of support via their vote share in the previous election(s). In most election settings, however, it is parties that designate the candidates entitled to appear on the ballot under their party label, sometimes choosing them by mechanisms purely internal to the party's formal organization, sometimes by "primary" elections with voting restricted to party members or those registered as party affiliates, sometimes in so-called open primaries where voters can choose the party primary in which they participate. Thus, it is often sensible to talk not about *the* electoral rules but about *electoral rules for a general election*, on the one hand, and *electoral rules for party nominations*, on the other, especially if different parties handle nominations in different ways.

Ballot Type

There are three main ways to distinguish types of ballots: (1) ballot complexity, (2) number of rounds of balloting, and (3) types of alternatives.

Ballot complexity refers to the kind of information that voters are required to provide. The simplest ballot is one where voters mark an X for some prespecified number of alternatives, or for *up to* some prespecified number of alternatives—for example, *approval voting*, where every voter may indicate that up to M alternatives are "satisfactory," with the M alternatives receiving the most "approval" votes being the ones that win; or *limited voting*, where each voter has a fixed number, k, of X ballots to cast, where k is less than the number of seats to be filled, M; and the most common case where voters have but a single X to cast, for example, *plurality voting* in a single-seat district. Another important type of X ballot is pure *list PR*, where there is a list prepared by each party, and a certain number of the top candidates on each list are elected, with that number determined by the proportion of (viable) votes cast for that party.

More complex ballots require voters to rank order alternatives, for example, the *Borda count* or *the single transferable vote*, or to assign cardinal numbers to alternatives, for example, *cumulative voting*. If one neglects the possibility of ties, one may think of the Borda method as involving each voter ranking all candidates and giving $C - k$ votes to the candidate the voter puts in kth place, where C is the number of candidates in the contest. For each candidate, their *Borda score* is the sum (over all voters) of these Borda counts, with the candidate with the highest Borda score to be elected. Under single transferable vote, voters also rank order the candidates. If there are M seats to be filled, any candidate who receives at least a *Droop quota* of votes ($=E/[M+1]$, where E is the size of the actual electorate) is elected, exactly $E/(M+1)$ of the ballots in which that candidate is at the top of the preference rankings are removed from further consideration, and the votes on the remaining ballots on which that candidate is at the top of the preference ranking are reallocated to the next highest ranked (still eligible) candidate on that ballot. If that reallocation now gives some additional candidate a Droop quota, then that candidate is elected, and we continue in this fashion as long as we can. If there are still unfilled seats, the candidate with the fewest first-place votes is dropped from eligibility, and his or her votes are reallocated to the next highest ranked (still eligible) candidate on the ballots of those who had the dropped candidate at the top of their preference ranking. This process continues until all seats are filled. (If there are some voters who do not rank sufficiently many candidates, their ballots may never come into play, and so sometimes it may be necessary to elect the last candidate[s] with less than a Droop quota.) There are many variants of cumulative voting, with perhaps the most common involving the requirement that all components of the allocation vector that must sum to the number of votes each voter is entitled to cast must be integers.

The two most common options for *numbers of rounds of balloting* are one-round and two-round ballots, but the exact number of rounds required in some forms of election (e.g., multiround sequential elimination elections) can vary with the distribution of votes among the various candidates. In a multiround sequential elimination system, there is some rule for dropping candidates at each round if no candidate has received a majority of first-place preferences among the set of still-viable candidates, for example, dropping the candidate with the fewest first-place votes from eligibility and then revoting among the remaining candidates.

Type of alternatives is meant to distinguish among the situations where voters choose among party lists versus voting for individual candidates versus some mix of the two—for example, *open list PR systems* or *mixed electoral systems*. In pure list PR, only parties are objects of choice, and the parties determine and rank their own candidates so that a party that wins *r* seats will elect the top *r* candidates on its list of candidates; in open list PR, in contrast, voters may also affect by their vote choices (even if not fully determine) which individual candidates from a given party will be elected. A *mixed electoral system* is, technically, simply one in which the electoral rule is not constant across all constituencies, but the term is more commonly used to refer to electoral systems that include both constituencies in which voters vote for a single candidate and those in which candidates are elected by some form of proportional representation (see Matthew Shugart & Martin Wattenberg, 2001).

But a further distinction is necessary as to what information about the candidate's party affiliation voters have available on the ballot. Even where voters are choosing among candidates rather than parties, voters usually know the party affiliation of each candidate or whether a candidate is running without a party attachment (i.e., as an "independent"). Nonpartisan elections, that is, elections without party labels for *any* of the candidates on the ballot, are relatively rare outside the United States, but in the United States such elections are the most common mechanism for electing local officials. Such nonpartisan elections are usually single-round, plurality-based elections, but there is a two-round variant called a *blanket primary*, used in a few U.S. states, where the top two vote getters enter into a second-round runoff if no candidate receives an actual majority of the votes cast. However, nonpartisanship often is merely a form of disguised partisanship, with the party affiliations of nominees being an open secret to knowledgeable voters.

Constituencies

Worldwide, most elections take place within geographically defined constituencies. As Bernard Grofman and Lisa Handley (2008) note,

> There are two kinds of exceptions to districted elections, the at-large election and the communal roll. In *at-large elections*, the entire polity is used as the district and thus there is no need ever to redraw constituency boundaries. Netherlands and Israel, for example, elect their national parliaments using List PR from the nation as a whole; while in the United States, the majority of cities elect city council representatives city-wide. In *communal rolls*, the fundamental basis of representation is non-geographic: choices are made from candidacies drawn from the members of a given race or ethnicity or religion. Usually, but not always, only members of a given community will be eligible to vote for representatives from that community. (p. 6, Footnote 3)

But, of course, even at-large elections, when they take place below the national level, are thought of as geographically based, and in many cases communal rolls may also be geographically based.

With respect to districted constituencies, a key distinction is between single-member districts (SMDs) and multimember districts (MMDs). The specification of constituency boundaries, called *redistricting* in the United States and *boundary delimitation* in much of the rest of the English-speaking world, is an important topic from both a legal and a theoretical point of view. For example, rules about the degree of population equality required across constituencies can be instrumental in permitting or preventing malapportionment, which, whether deliberate or unintended, can have substantial consequences for the translation of votes into seats and the representation of groups that differ in their geographical locations and degree of geographic concentration, such as blocs of ethnic voters or party supporters.

Election Timing

In terms of election timing there are two important distinctions. The first is between elections that take place at predetermined intervals versus elections that may be "called" by those in power or that result from the failure of a vote of confidence in the parliament. When election dates are not predetermined, there is scope for strategy in the timing of elections, and parties within a governing multiparty coalition may have credible threats to withdraw from the cabinet so as to force a new election. Such threat potential can create policy uncertainties and ultimately destabilize governments.

The second distinction, involving election timing, is about the degree to which elections of various types and at various levels of government are held concurrently. Generally speaking, when elections are *not* held concurrently, this increases the likelihood that different policies will be pursued at different levels of government and that different parties will be in power for different offices.

Ballot Aggregation

The process by which ballots are turned into outcomes is referred to as ballot aggregation in the social choice literature.

The most frequently made distinction among ballot aggregation methods is between *majoritarian/plurality methods* and *methods of representation* intended to allow unified blocs of voters of a certain size that do not constitute an overall majority to nonetheless gain representation. The best known form of minoritarian representation is *proportional representation.* The *proportionality principle* refers to the idea that seat shares in a legislature should reflect vote shares.

The distinction between SMDs and MMDs is linked to this distinction, in that proportional representation methods can only be used in MMDs; however, use of MMDs does not imply the use of proportional representation, since there can be plurality voting in MMDs (sometimes called *plurality bloc voting*), where each voter has M (single) votes to cast and where the M candidates with the most votes are elected, or *limited voting* in MMDs, which, for $k > 1$, is a *semiproportional* system (see below).

The standard way to conceptualize the continuum anchored by majoritarian/plurality methods, at one end, and proportional representation methods, at the other end, is in terms of the *threshold of exclusion*, which Douglas Rae, Victor Hanby, and John Loosemore define as the largest vote share a party can receive and still be denied any seats in the district (see Rae, 1967/1971). The threshold of exclusion involves a hypothetical worst-case scenario in terms of the distribution of votes among the other parties that will minimize the given party's seat share. For M seat districts, in pure PR systems, the threshold of exclusion is either exactly or approximately $1/(M + 1)$. In other words, if it coordinates its votes, a group that has only a $1/(M + 1)$th share of the district's population can guarantee to elect a candidate of its choice. As M gets larger, smaller and smaller minorities become a large enough share of the electorate to win a seat for a candidate of their choice. The minimum feasible threshold of exclusion, which occurs for PR systems with the whole polity as the single constituency, is $1/(S + 1)$, where S is the number of seats in the parliament. At the other end of the continuum, for simple plurality and majority systems, the threshold of exclusion is (infinitesimally under) 1/2, regardless of M. Intermediate cases arise for limited voting, where voters have k (single) votes to cast, and there are M seats to be filled ($k < M$), with the M highest vote getters elected. Here, the threshold of exclusion is $k/(k + M)$. The closer k is to M, the less proportional is the limited voting rule. The threshold of exclusion for SNTV, the single nontransferable vote, which is another name for limited voting in M seat districts ($M > 1$) where voters have but a single vote to cast, is identical to that of the d'Hondt form of list PR, namely $1/(M + 1)$.

Within any given country, an m-seat district can be expected to have roughly m times the population of a single-seat district. Thus, if one is interested in the actual number of voters who must change their mind to affect election outcomes, then threshold of exclusion values need to be adjusted to take population differences into account across constituencies of different sizes.

Paralleling the threshold of exclusion, there is also a *threshold of representation*, which is the smallest vote share a party can receive and still win a seat in the district (see Rae, 1967/1971). The threshold of representation involves a hypothetical best case scenario in terms of the distribution of votes among the other parties that will maximize the given party's seat share.

The thresholds of exclusion and representation are theoretical concepts. In any given electoral system, the actual vote share that a party will need to be assured of to gain representation will often be much less than the theoretical maximum. Relatedly, as Rein Taagepera (personal communication, June 4, 2004) observes, "The same electoral rules can lead to vastly different disproportionality, even in the same country and even in consecutive elections." To deal with this problem, scholars commonly calculate *empirical* indices of disproportionality.

For *partisan elections*, the two most common measures of overall proportionality are the Loosemore-Hanby Index of Distortion (see Rae, 1967/1971),

$$D = \frac{1}{2}\sum|v_i - s_i|,$$

with v_i being the vote share of the ith party and s_i the seat share of that same party; and the Gallagher Index (Michael Gallagher, 1991),

$$\mathrm{Gh} = \left[\frac{1}{2}\sum(V_i - S_j)^2\right]^{0.5}.$$

For two-party competition, another approach to measuring disproportionality empirically is in terms of what is called the *swing ratio*. Edward Tufte (1973) proposed that, in two-party legislative competition, a party can expect to receive a share of seats such that

$$\frac{\log s}{1-s} = \frac{k \log v}{1-v} + \varepsilon.$$

Here, the closer k is to 1, the closer we are, empirically, to proportionality.

While the PR versus plurality/majority distinction is certainly important, there are equally important distinctions among the members of each of the two groupings.

Among PR methods, the most important threefold distinction is between list PR methods and rules that are based on preference rankings, most notably the single transferable vote (STV, aka the Hare system), on the one hand, and those that involve point schemes, most notably the simple cumulative vote (CV), where each voter has M votes to be divided (in whole integers) among candidates in any way that the voter wishes, on the other hand. While CV, like SNTV, is sometimes called semiproportional since it requires voter coordination to ensure proportionality, this is a quantitative rather than a qualitative difference between them and, say, list PR methods, and so both are treated under the general rubric of PR. The intermediate category of semiproportional systems is reserved for limited voting, with $1 < k < M$.

There are also important distinctions among list PR systems. This entry has already mentioned the issue of whether the system allows voters to choose among candidates rather than parties. A second, rather more technical, distinction is between methods that can be characterized as *divisor methods* and methods that cannot be so characterized, most notably the *greatest-remainder* method. Divisor methods are methods where one can determine outcomes by taking the party vote shares p_i, and dividing each in turn by a particular vector of divisors $(d_1, d_2, \ldots, d_M)$ and then allocating seats to each party corresponding to their having one of the M highest quotients from among the $n \times M$ member set $\{p_i/d_j\}$. For example, d'Hondt (also sometimes known as the Hagenbach-Bischoff method) uses the vector of divisors $(1, 2, 3, 4, \ldots, n)$, while Sainte-Laguë uses $(1, 3, 5, 7, \ldots, 2n - 1)$, and the modified Sainte-Laguë uses $(1.4, 3, 5, 7, \ldots, 2n - 1)$.

For example, if there were three parties of sizes 0.40, 0.35, and 0.25, respectively, and six seats to be filled, under d'Hondt one would look at the set of numbers {0.40, 0.35, 0.25; 0.20, 0.175, 0.125; 0.133, 0.117, 0.083; etc.} and find the six highest, which would result in three seats to the largest parties (from quotients of 0.40, 0.20, and 0.133), two seats to the second largest party (from quotients of 0.35 and 0.175), and one seat to the smallest party (from a quotient of 0.25). Since the quotient of 0.40/3 (=0.133) is higher than 0.25/2 (=0.125), the sixth seat goes to the largest party rather than the smallest party. To calculate the greatest-remainder results, one would look at $p_i \times M$ and first allocate whole quotients to the parties and then allocate the remaining seats in order of the magnitude of the remainder. For the same six-seat example with three parties of sizes 0.40, 0.35, and 0.25, one gets 2.4, 2.1, and 1.5, respectively. Since the greatest remainder is 0.5, each party ends up getting two seats, despite the disparities in size among the three parties.

Note that even though both d'Hondt and the greatest remainder are PR systems, they do not always yield identical allocations. In general, although most often the allocations under the two rules will be the same, sometimes d'Hondt will be more favorable to the largest/larger parties than will the greatest remainder. But not all divisor rules produce the same results, either. In particular, Sainte-Laguë is likely to be in between the other two rules in terms of favoring larger over smaller parties, while the modified Sainte-Laguë was specifically designed to be more favorable to moderately sized parties than the unmodified Sainte-Laguë.

Mathematically, divisor methods for allocating seats to parties on the basis of party vote shares are identical to divisor methods for allocating seats to geographic units on the basis of the unit's share of the total population. The same method may go under different names in the two different contexts. Thus, the d'Hondt method was actually first invented by Thomas Jefferson to partition seats in the U.S. House of Representatives among the states. Similarly, the Sainte Laguë method is identical to an apportionment method devised by the American legislator Daniel Webster. Five divisor methods have particular centrality because they each correspond to a particular way of rounding values to create integer allocations of seats to the various parties that sum up to M. The Adams method corresponds to rounding up; the Jefferson method (d'Hondt) corresponds to rounding down; the Webster method (Sainte-Laguë) corresponds to rounding up if the remainder is greater than 0.5; the Dean method corresponds to taking the harmonic mean; and the Hill-Huntington method corresponds to taking the geometric mean and is the method that is currently in use in the U.S. House of Representatives to determine how many seats each state will get (subject to the constraint that each state must have at least one seat in the House). There are also alternative ways to define divisor methods, for example, in terms of quotient maximization functions. It should also be noted that the greatest-remainder method is also called the Hamilton method because Alexander Hamilton proposed it as way of allocating seats to the states.

While the distinctions among types of ballot aggregation identified above are the ones most often studied—generating what Rein Taagepera (2007) refers to as the set of *simple electoral systems*—they are only the tip of the iceberg. The devil is in the details, and there are many other features of electoral rules that can affect vote aggregation. For example, there may be rules that allow parties to combine their vote share by entering into preelection alliances; and/or there may be specific national, regional, or district vote share thresholds or rules for geographic distribution that must be surmounted before a party can receive any seats; and/or there may be complex mixtures of electoral rules, with different types of rules applying in different constituencies or in multiple tiers of representation involving different levels of proportionality and "topping up" across the various tiers.

Electoral Rule Effects

There are five major theoretical perspectives on how to estimate the likely consequences of electoral rules or changes in electoral rules: (1) social choice theory, (2) mainstream empirical research, (3) rational choice and game-theoretic models, (4) the social physics perspective, and (5) the embedded systems approach.

Social choice approaches are those that (a) in the spirit of Kenneth Arrow (1951/1963) study the normative properties of voting rules, for example, responsiveness to changes in underlying voter preferences, the likelihood of choosing a candidate who can receive a majority over all other candidates in a head-on-head contest (known in this literature as *Condorcet efficiency*), and/or seek to characterize voting rules in axiomatic form; (b) are concerned about the structure of individual preferences, for example, the extent to which preferences have a particular form of unidimensionality known as *single peakedness* or satisfy other domain restrictions; and/or (c) for a fixed set of voter preferences, assess the likelihood that different voting methods will yield the same outcome.

Mainstream empirical research seeks to measure the effects of particular electoral rules, cross-nationally or across different units in the same polity, by techniques such as regressing an outcome variable—for example, (effective) number of political parties or measures of proportionality of seats–votes relationships—against electoral system features and some set of control variables (see, e.g., Arend Lijphart, 1994). It is characterized by large-scale data analysis of both survey and aggregate data.

Rational choice modeling emphasizes strategic behavior on the part of voters and candidates/parties and makes extensive use of game theory. Customarily, results are in the form of theorems about how electoral system effects, in equilibrium, are determined by the incentives that different rules provide for the behavior of voters and parties/candidates under different assumptions about the motivations ascribed to voters and the utility functions (office seeking, policy seeking, or some combination thereof) that are ascribed to parties/candidates (see, e.g., Gary Cox, 1997). While the formal properties of electoral rules determine how "inputs," that is, completed ballots, will be converted into electoral outcomes, it is the structure of electoral incentives that helps determine both what options will actually be available to the voters on the ballot and how voters will decide among those options. In particular, when there is *strategic voting*, such that voters do not always support the candidate/party they most prefer if they do not believe that that candidate/party has a realistic chance to be elected, then great care must be taken in projecting outcomes from underlying preferences and the nature of the voting rule.

The s*ocial physics approach*, associated with the work of Rein Taagepera (see Taagepera & Matthew Shugart, 1989; also especially Taagepera, 2007), is inspired by the principles of statistical thermodynamics in physics. It does not attempt to predict the effects of electoral rules in individual political units but seeks instead to precisely model effects on average. Also, it (a) makes use of only a handful of key variables; (b) uses functional forms, which must yield results consistent with the boundary conditions determining the range of feasible outcomes; and (c) requires that left-hand (dependent) and right-hand (independent) variables be stated in a fashion that yields dimensional comparability.

The term *embedded systems* was introduced in Grofman (1999) and used in subsequent work by Grofman and coauthors, but many others use the same approach without calling it by that name. The hallmark of the embedded systems approach to electoral systems is the notion that electoral system effects can be fully understood only within the context of the overall constitutional, social, and party systems in which they are embedded. Of special concern are (a) seeing how similar systems can yield different outcomes in different contexts; (b) the need for care in attributing causality to electoral system effects when the choice of electoral rules may be endogenously determined, which leads to an interest in experiments on voting rules and "natural experiments" such as those that occur when different methods are used within the same polity in different geographic areas or at different time points; (c) attention to how the seemingly trivial differences in electoral rules, for example, different rules for nominating candidates, can have major consequences; and (d) attention to the fact that it will take time until voters, candidates, and parties come to understand how a given set of electoral rules, embedded as they are within a particular political context, will affect outcomes.

While social choice theory has, historically, largely limited itself to studying the properties of voting rules in the abstract, many political scientists and economists have been more interested in who/what is being affected by the choice of rules. Scholars have looked at the implications/consequences of electoral rules for voters, parties, and interest groups as well as at general effects on governance and effects on specific policy outputs.

For voters, there is an extensive literature on how electoral systems may affect representation of geographic areas and of racial/ethnic/social groupings, and voter turnout, as well as on spillover effects, such as on voter perceptions of transparency and voter efficacy. For parties, there is work on the consequence of electoral rules for party proliferation, party ideological dispersal, strength of party organization, party discipline in the legislature, party coalitional strategies, and of course, the relationships between party vote shares and party seat shares in terms of the extent of proportionality and of partisan bias. For governance, important questions have to do with the link between electoral rules and issues such as the incentives for corruption, the incentives for localism and pork-barrel politics, the durability of parliamentary governments, and for ethnically divided societies, the links between electoral rules and civic unrest and the strength of secessionist movements. For specific policy arenas, hypotheses have been proposed about the (independent) effects of electoral rules on the size of the welfare system and the degree of economic inequality (see, e.g., Torsten Persson & Guido Tabellini, 2003).

The electoral systems literature is now so vast that no single essay can possibly do justice to it. Readers are reminded of two points: First, one must be attentive to issues of endogeneity and interaction effects. Voting rules do not exist in a vacuum; on the one hand, they are determined by past history and maintained only in some type of political equilibrium, and on the other, their consequences are shaped by the present political context in which they are embedded even though they also affect that context. Second, and relatedly, to understand electoral system effects, one must understand both the formal properties of electoral rules (including prosaic facts such as the maximum number of parties elected in a constituency being capped by the number of seats that are up for election), on the one hand, and the (short-run and long-run) incentives for voters, candidates, parties, and interest groups created by those rules, on the other.

Bernard Grofman
University of California, Irvine
Irvine, California, United States

See also Election by Lot; Elections, Primary; Election Research; Elections, Volatility; Electoral Behavior; Electoral Geography; Electoral Turnout

Further Readings

Arrow, K. J. (1963). *Social choice and individual values.* New York: Wiley.

Cox, G. (1997). *Making votes count: Strategic coordination in the world's electoral systems.* New York: Cambridge University Press.

Gallagher, M. (1991). Proportionality, disproportionality and electoral systems. *Electoral Studies, 10,* 33–51.

Grofman, B. (1999). SNTV, STV, and single-member district systems: Theoretical comparisons and contrasts. In B. Grofman, S.-C. Lee, E. Winckler, & B. Woodall (Eds.), *Elections in Japan, Korea and Taiwan under the single non-transferable vote: The comparative study of an embedded institution* (pp. 317–333). Ann Arbor: University of Michigan Press.

Handley, L., & Grofman, B. (Eds.). (2008). *Redistricting in comparative perspective.* Oxford, UK: Oxford University Press.

Lijphart, A. (1994). *Electoral systems and party systems: A study of twenty-seven democracies, 1945–1990.* Oxford, UK: Oxford University Press.

Persson, T., & Tabellini, G. (2003). *Economic effects of constitutions.* Cambridge: MIT Press.

Rae, D. (1971). *Political consequences of electoral laws* (2nd ed.). New Haven, CT: Yale University Press.

Shugart, M., & Wattenberg, M. (2001). *Mixed-member electoral systems: The best of both worlds?* New York: Oxford University Press.

Taagepera, R. (2007). *Predicting party sizes: The logic of simple electoral systems.* New York: Oxford University Press.

Taagepera, R., & Shugart, M. (1989). *Seats and votes: The effects and determinants of electoral systems.* New Haven, CT: Yale University Press.

Tufte, E. (1973). The relationship between seats and votes in two-party systems. *The American Political Science Review, 67*(2), 540–554.

ELECTORAL TURNOUT

Electoral turnout is called an "aggregate-level" concept because it only makes sense to talk of turnout when talking of aggregates of individuals. Yet because turnout derives from cumulating individual-level decisions, one must start with an understanding of why people vote—something about which there is little agreement in the scholarly literature. Proposed theories have been based on rational, expressive, mobilized, habitual, dutiful, altruistic, and even genetic explanations; see, for example, Benny Geys (2006) and James Fowler and Christopher Dawes (2008). From a rational choice perspective, the decision to vote is conceptualized as the result of a cost–benefit "calculus of voting" in which difficulties of becoming informed and getting to the polling booth, and so on, are compared with the benefits a voter expects if his or her preferred party is enabled to implement promised policies. The calculus must take account of the probability that one vote will affect an election's outcome. In a large electorate, this probability is vanishingly small, so a truly rational voter would "free ride" on the efforts of others. The fact that in reality many citizens do vote has come to be called the "paradox of voting."

Electoral Participation as Social Behavior

The recognition of the paradox of voting led some scholars to reinterpret benefits and costs in terms that are unrelated to outcomes of elections by assuming that people are motivated by social, ethical, or expressive considerations rather than narrow self-interest. In brief, ethical explanations claim that citizens vote out of a sense of civic duty, mobilization theories argue that people are motivated by political campaigns, while expressive explanations suggest that individuals are affirming their (group) identity. Proponents of these theories have not, however, explained why, if motivations are unrelated to election outcomes, people would be more likely to vote when their vote has more impact on the outcome (e.g., in a close race) or when more is at stake (see below).

An approach with more promise of overcoming the paradox of voting recognizes that voting benefits more people than just the individual voter and sees voters as social beings motivated in part—especially if they are of an altruistic disposition—by their obligations to groups with which they identify and responsive to the expectations and mobilizing efforts of members of those groups. This approach has gained prominence with recent research demonstrating that turnout is subject to socialization, learning, and habit formation, as shown by Eric Plutzer (2002). In addition, the value of this approach is in demonstrating that most voters acquire a habit of voting over time and that contextual factors such as the closeness of the race have a stronger impact on voters who have not yet acquired the habit of voting, as demonstrated by Mark Franklin (2004).

The fact that people differ in terms of whether their behavior is based on habit suggests that much earlier research into individual-level explanations of turnout variations, demonstrating the influence of sociodemographic factors (such as education, income, or age) and social factors (such as home ownership or church attendance), were focusing on factors that facilitated this learning process. A reconceptualization of the decision to vote into long-term (habitual) and short-term (variable) behavior shows promise of clarifying much that is currently confusing in the turnout literature.

Short- and Long-Term Forces in Turnout Change

The same distinction between short- and long-term forces helps explain a puzzle that has not received much attention. Turnout differences between countries are considerable. In some countries (e.g., Australia, Belgium, or Malta), virtually everyone votes. In other countries (Switzerland or the United States), turnout rarely reaches 50%. These differences are huge, yet in established democracies the turnout seldom varies by as much as 5% between successive elections of the same type (presidential elections in the United States, parliamentary elections elsewhere). Only over a longer period (20 years or more) do some countries see marked changes in their turnout levels. This puzzle is resolved if one understands that most people in established democracies at any given time have acquired established habits of voting or not voting—habits that condition the level of turnout according to the ratio of habitual voters to habitual nonvoters. An electorate consisting mainly of habitual voters will regularly produce a high turnout, whereas an electorate with more nonvoters will produce lower turnout, and these levels of turnout will be quite stable in the short run. Any long-term change in turnout levels must come from new voters whose behavior is not yet habitual. If the proportion of new voters who acquire the habit of voting changes and remains at a different level (higher or lower), then long-term change will arise from natural processes of generational replacement. Such processes are slow but inexorable because an electorate will be entirely replaced with new voters over the course of about 50 or 60 years.

What would lead new voters to acquire the habit of voting at a different rate than their elders did? To answer this question, one needs to turn to explanations of turnout variations that focus on things about elections rather than things about individuals.

Election-Level Explanations

Changes in the level of turnout between countries and over time are not due to changes in individual-level characteristics. Increases in the proportion of university-educated or high-income individuals (two characteristics strongly associated with the acquisition of habitual voting) do not result in rising turnout. The United States and Switzerland are both rich countries with very high levels of education, yet their turnout is lowest of all established democracies; and in both countries, turnout fell even as wealth and education increased. Moreover, countries that see high turnout do not have markedly different distributions of these and other social characteristics in their electorates. So there must be something else that can motivate everyone—even the poorest and least educated—to vote in certain circumstances, quite overriding the influence of individual-level factors—something that must be weaker or even absent in other circumstances.

Scholars agree that this "something" relates to what is at stake in an election. In countries where much is at stake, everybody votes. Where the stakes are lower, so is turnout. Research has pointed to several factors that raise the stakes of an election—primarily, the institutional and competitive context (see Andre Blais, 2006; Benny Geys, 2006). To indicate the nature and importance of the more potent contextual factors, Table 1 displays their effects in 22 established democracies between 1945 and 1999. The right-hand column shows the maximum extent of turnout differences found to be associated with each factor, comparing elections with the lowest stakes in terms of that factor to elections with the highest stakes.

These influences are ordered in terms of their importance in the long-term perspective (i.e., after allowing time for generational replacement to reveal their full effects), with the most influential characteristic being "policy responsiveness." This influence is emblematic because it relates so clearly to the policy stakes of an election: Some political systems have separated powers or other obstacles that stand in the way of a majority enacting its policies, reducing the extent to which voters can easily estimate the policy consequences of an election. "Electorate size" is almost as important, picking up additional aspects of policy responsiveness since larger countries are more often subdivided by means of federal or quasi-federal arrangements. Electorate size also dilutes the power of each individual vote. "Largest party size" is a measure of legislative power. A party that reaches the 50% threshold can govern alone; one that comes close can dominate any coalition of which it

Table 1 Maximum Effect of Differences in the Stakes of an Election

Influence	*Range of Values*	*Maximum Turnout Difference Due to This Influence (%)*
Policy responsiveness	0–3 units	16.0
Electorate size	1–280 million	12.3
Compulsory voting	0, 1 = yes	11.3
Largest party size	33–50%[a]	7.0
Mean margin across districts	0–10%	5.7 [b]
Margin of victory nationwide	0–10%	5.3 [b]
Party cohesiveness	0, 1 = yes	4.2
Sum of differences		56.5 [b]

Source: Adapted from Table 5.2 in Franklin, M. N. (2004). *Voter turnout and the dynamics of electoral competition in established democracies since 1945*. New York: Cambridge University Press.

a. Maximum limited to 50%, the threshold for majority status.

b. Margin and mean margin operate in different countries, so one is zero when the other is not.

forms a part, yielding confidence that it will enact its policy promises (other factors permitting). One obvious prerequisite is "party cohesiveness" at the foot of the table—that is, the extent to which parties enjoy disciplined support in the country's legislature. Two alternative measures of margin of victory are shown, depending on whether votes are cast in districts where the "winner takes all" or seats are assigned proportionately to votes. Both measures tap the extent to which an election is a close one, in which every vote counts. "Compulsory voting" has a rather different influence, increasing the stakes of an election by threatening sanctions (real or imagined) for failure to vote. In practice, this influence has been shown to have most effect where turnout would otherwise be lowest, effectively trumping all other influences on turnout (Arend Lijphart, 1997).

The final coefficient at the bottom of the table shows the long-term difference in turnout that might be observed empirically between a hypothetical election in which all stakes are at their maximum and one where all stakes are at their minimum. It is something of an underestimate because several small influences have been omitted for simplicity. However, no established democracy in practice has elections characterized by minimum possible stakes, so the hypothetical total difference of 56% is not an unrealistic maximum.

In new or transitional democracies, greater swings in turnout can occur than generally seen in established democracies—presumably because in new democracies no one has yet had time to acquire habitual behavior patterns. Indeed, the more volatile turnout seen in such countries serves to reinforce our conclusions about the role of habitual behavior in turnout change. Nevertheless, little is known about how the mechanics of turnout change in established democracies play out in new democracies (Andre Blais & Agnieska Dobrzynska, 1998; Pippa Norris, 2002). So the study of new and transitional democracies remains an important avenue for future research.

Further research is also needed to test competing explanations of individual-level behavior. At the aggregate level, more research is needed on new democracies as well as research connecting the aggregate context of elections to individual voter behavior. The methodological advances being made in time series and multilevel analysis should help answer the remaining puzzles.

Mark N. Franklin and Carolien Van Ham
European University Institute
San Domenico di Fiesole, Italy

See also Civic Participation; Democratization; Electoral Behavior; Electoral Systems; Multilevel Analysis; Participation; Rational Choice; Time-Series Analysis

Further Readings

Blais, A. (2006). What affects voter turnout? *Annual Review of Political Science, 9,* 111–125.

Blais, A., & Dobrzynska, A. (1998). Turnout in electoral democracies. *European Journal of Political Research, 33,* 239–261.

Fowler, J. H., & Dawes, C. T. (2008). Two genes predict voter turnout. *Journal of Politics, 70*(3), 579–594.

Franklin, M. N. (2004). *Voter turnout and the dynamics of electoral competition in established democracies since 1945*. New York: Cambridge University Press.

Geys, B. (2006). "Rational" theories of voter turnout: A review. *Political Studies Review, 4,* 16–35.

Lijphart, A. (1997). Unequal participation: Democracy's unresolved dilemma. *American Political Science Review, 91,* 1–14.

Norris, P. (2002). *Electoral engineering: Voting rules and political behavior*. New York: Cambridge University Press.

Plutzer, E. (2002). Becoming a habitual voter: Inertia, resources, and growth in young adulthood. *American Political Science Review, 96,* 41–56.

ELITES

Elites are small groups of persons who exercise disproportionate power and influence in social domains. It is customary to distinguish between *political elites*, whose locations in powerful institutions, organizations, and movements enable them to shape or influence political outcomes, often decisively, and *cultural elites*, who enjoy a high status and influence in nonpolitical spheres such as arts and letters, philanthropy, professions, and civic associations. At the national level, political elites number only a few thousand persons in all but the largest countries, whereas the makeup of cultural elites is more indeterminate and turns on the nonpolitical spheres regarded as consequential in a society. Taking their lead from Gaetano Mosca (1856–1941), Vilfredo Pareto (1848–1923), and Robert Michels (1876–1936), political scientists focus on political elites (simply elites hereafter), investigating how their actions and beliefs, social profiles, and overall configurations affect political regimes and policies. Below, various aspects of elite formation, autonomy, structure, transformation, and their analysis in political science are discussed.

Mosca, Pareto, and Michels portrayed elite formation as inescapable in modern societies and as imposing limits on what is possible in politics. They maintained, for example, that genuine democratic systems are impossible because there are always self-interested elites who out-organize and outwit the *demos*. The most that can be hoped for, in their view, is a relatively liberal but still quite unequal order led by capable, cooperative, and enlightened elites. But, they noted, elites in most societies and times fall well short of these attributes, so politics usually involve fierce power struggles between ambitious, blinkered, and insecure elites.

The research literature on elites is large. Especially prominent are studies of parliamentary and cabinet-level politicians; high civil servants; and the owners, CEOs, and board members of major business firms. Studies of military, trade union, mass media, pressure group, religious, and social movement elites are also numerous. There is, in addition, a substantial body of historical research uncovering the makeup and role of elites in pivotal events and processes of economic, political, and social change. Researchers identify elites in terms of who holds the most important formal positions in a society, who has pronounced reputations for political power and influence, or who participates in making key decisions and policies. Each identification method tends to produce a distinctive elite configuration. Concentrating on positions highlights the differentiation of elites by sectors (politics, business, government administration, etc.); focusing on reputations often identifies a cohesive and small "power elite" sitting astride diverse sectors; and studying participation in decisions and policies most often yields an array of specialized elite groups. To overcome these identification peculiarities, researchers frequently use more than one method and combine the results.

Once elites are identified, how their positions overlap and interlock can be analyzed, and data on their social background, education, and careers can be compiled from public registers and questionnaires; also, elite persons' opinions and perceptions, as well as their informal interactions and networks, may be solicited through interviews. A perennial question is the degree to which collective elite profiles mirror those of nonelite publics and, thus, the extent to which elites can be supposed to reflect and represent such publics. Other important questions include the extent to which elites are autonomous actors, how their compositions and behavior vary among societies, and how elites change, gradually or suddenly, in these and other respects.

Elites and Nonelites

Research shows that in their social background, education, and occupations, elites are almost always more privileged than cross sections of nonelite populations. They come prevailingly from well-off families, some of whose members may have held elite positions in preceding generations. The frequency with which they hold university degrees—often from "elite" institutions—far exceeds the distribution of such education among nonelite publics. Elites also come disproportionately from high-status occupations—lawyers, teachers, and public or private sector managers. There are exceptions, of course. Significant proportions of civil service elites have career graphs that begin in lowly positions and involve long climbs to the top. Self-made entrepreneurs are certainly not unknown among business elites. Historically, the elites of trade unions and various social movements exhibited a modest background and education, although today in industrialized countries, they are predominantly middle class in background and university educated. Elite-level politicians are increasingly involved in full-time careers in or close to politics, often starting as student political leaders and then serving on leading politicians' staffs or holding paid positions in parties, or they come from politically relevant careers in journalism, public relations, and policy think tanks. Historically, elites have consisted almost entirely of men, and men continue to outnumber women greatly in most elite sectors. In multi-ethnic or multiracial societies, elite persons usually belong to the largest or otherwise dominant ethnic or racial population segment. Finally, in average age, elites tend to be significantly older than nonelites.

Research shows, however, that these differences between elites and nonelites are gradually being reduced. For example, research on the social, educational, and occupational profiles of parliamentary elites in 11 European countries from the middle of the 19th century to the present reveals a long trend toward less exclusive and privileged profiles. Changes in the gender makeups of elites are also now quite evident. It is possible, moreover, that the preponderance of bureaucratic and service work in many contemporary societies is intermingling elites and nonelites in important ways. Both perform essentially similar nonmanual work tasks and not infrequently rub shoulders in offices. Because elites now more frequently ascend to their positions from nonelite origins than in the rigidly stratified societies of earlier times, not a few see themselves as one of a kind with nonelites, among whom they have intimate personal associates and for whom they have considerable empathy. Such closer and more empathetic ties may dispose elites toward actions that better reflect and represent nonelite desires and interests. There is, furthermore, a debate about the degree to which the backgrounds of elites translate into opinions and actions that disregard nonelites. Some scholars argue that there is at most a weak linkage between elite backgrounds and elite opinions and actions. Faced with pressing and constantly changing organizational and political imperatives, elites hold views and take actions that resonate poorly, if at all, with their life histories. Elite backgrounds may reveal more about the distribution of opportunities in a society than why its elites think and act as they do.

Elite Autonomy

Elites seldom enjoy complete autonomy. To carry out major initiatives and perpetuate their holds on power, elites need nonelite support. To win it, elites frame appeals that accord, at least roughly, with nonelite interests and political orientations. It is reasonable to suppose that nonelite interests and orientations are not created by elites but derive from combinations of human nature, life circumstances, and beliefs, although elites often activate or muffle nonelite interests and orientations through well-couched appeals. Failure to win nonelite support frequently shortens elite tenures or undermines elite effectiveness. Still, elites normally have a range of choices that are decisive for political outcomes. In political crises, for example, opposing elite camps and factions may choose to enter into accommodations or persist in lethal oppositions. During transitions to democracy, ascendant factions may choose to embrace or merely feign democratic practices. In well-instituted democracies, elites may proffer outcomes they know to be impossible or resort to obfuscations that enable them to avoid responsibility for what happens.

How much autonomy elites have—and should have—are questions that have long dogged discussions of "democratic elitism." In his seminal book *Capitalism, Socialism and Democracy*, Joseph Schumpeter (1883–1950) held that democracy is simply a method by which voters select governing leaders and elites, who should then be left alone to get on with the business of governing. For Schumpeter, in other words, democracy combines autonomous governance by leaders and elites with time-limited mandates to govern given by the *demos* in periodic elections. However, many critics of democratic elitism claim that this too blithely assumes leaders and elites to be creative and responsible actors who can safely be entrusted with autonomy. Whether the autonomy of leaders and elites is increasing or decreasing in contemporary democracies is a major empirical question in political science, and its desirability is a bone of contention in normative democratic theory.

Variation in Elite Structure and Values

It is obvious in modern history and today's world that elites differ importantly in their structures and values. First, elites located in functionally differentiated societal sectors display different modes of organization and different aims. Much research on the careers, outlooks, and skills peculiar to elites in specific sectors, as well as the ease or difficulty with which elite persons traverse sector boundaries or hold interlocked positions in two or more sectors, assesses the extent to which elites are pluralistic or monolithic in their structures and values.

Second, there are efforts to capture different types of national elites when considered as wholes. These efforts risk being too simplistic, yet the comparative study of elites cannot escape the need for a broad typology. Typologies that have been put forth concentrate on the extent of structural integration and the extent of value consensus. Structural integration involves the relative inclusiveness of formal and informal networks of communication and influence among elite members. Value consensus involves their relative agreement about the worth of existing political institutions and a code of political behavior.

Comparing national elites along these dimensions, a disunited type and two types of united elites can be distinguished, although there are difficult borderline cases. In the disunited type, elite groups are clearly divided and separated from each other, they disagree fundamentally about the worth of existing institutions, and they adhere to no single code of behavior. This is still the most common type of elite in the modern world. It is usually associated with authoritarian regimes in which one elite group or camp has the upper hand and governs ruthlessly. A disunited elite may also be associated with illiberal democracies, in which competitive elections take place but are so ridden with fraudulent practices that few elites and citizens accept their outcomes as legitimate.

Disunited elites have originated principally in processes of national state formation, during which elites in one territory forcibly annex and subjugate other territories and elites. The resulting conflicts and hatreds have most often produced deeply disunited elites in new national states. There are many examples: the animosities and bitter memories that followed the formation of national states in Europe between the 16th and 19th centuries, the "man on horseback's" (caudillos) control of Latin American states from their formations early in the 19th century until long into the 20th century, nearly all African states under alternating military and civilian rule since their emergence in the 1950s and 1960s, and the creation of postcolonial states by dynastic ruling elites in the Middle East.

In one of the two types of united elites, by contrast, elite structure is sharply centralized in a party, theocratic, or ethno-nationalist state. There is a secular ideology, religious doctrine, or ethnonationalist creed that is defined by the uppermost political leaders and professed publicly by all or nearly all elite persons. Much less common than the disunited type of elite, examples of this "ideologically united" type are the Soviet elite from 1922 until the mid-1980s, Germany's elite during the Third Reich, the North Korean elite under Kim Il-sung and his son Kim Jong-il, China's elite during Mao Zedong's rule and after, the Cuban elite under Fidel Castro, the Viet Minh elite since gaining control of North Vietnam in 1954 and all of Vietnam in 1975, the Khmer Rouge elite in Cambodia (until 1978), and the Iranian elite from the time of Ayatollah Khomeini's ascendancy (1979) and, arguably, until the present. As these examples indicate, the ideologically united type of elite most often originates in a revolutionary or anticolonial war that culminates in the victory and dominance of an especially doctrinaire elite group.

Disunited and ideologically united elites are clearly incompatible with stable liberal democracies, whereas a second type of united elite—what might be labeled a "consensually united" elite—appears to be a precondition for them. In this second type, elite structure consists of formal and informal networks of influence and acquaintance that are most dense within the functionally differentiated sectors mentioned above. But these networks overlap and interlock to form web works and "inner" or "central" circles through which sector elites are tied together and obtain access to key political decision makers. As well, the members of this second type of united elite share a voluntary, mostly tacit consensus about political norms and practices, the hallmark of which is keeping political competition restrained and nonviolent. Factions recognize each other's right to be heard, they agree to disagree when decisions cannot be reached, they emphasize technical and procedural feasibilities rather than ultimate rights and wrongs, and they practice enough secrecy to have flexibility when bargaining and fashioning compromises on difficult policy issues.

Like ideologically united elites, consensually united elites are relative rarities in the modern world. They appear to originate in only two or three ways. One is through a deliberate, sudden, and basic "settlement" of costly and long-standing oppositions, as occurred between Tory and Whig elite camps in England's Glorious Revolution during 1688 to 1689 and between *Franquist* and anti-*Franquist* elite camps in Spain during 1977 to 1978. A second way has been through a long experience of relatively benign colonial home rule during which settler or "native" elites learn to practice cautious and restrained politics and eventually make a unifying effort to gain independence—the origin of consensually united elites in the former British colonies of Australia, Canada, India, Jamaica, New Zealand, and, not least, the United States. A third way has been the gradual convergence of camps in a disunited elite toward consensual unity when spreading prosperity and agreement about basic democratic rights and principles, which erodes the postures taken by radically dissident camps, as happened in France and Italy during the 1960s and 1970s and Germany and Japan during the 1970s and 1980s. It must be noted, however, that basic settlements between opposed elites are highly contingent and have been historically relatively rare occurrences. The further consolidation and formation of elite consensus in the more recent "Third Wave" (Huntington) democracies remains to be seen.

Elite Circulation and Transformation

How elites change in their composition and behavior is a major question. Mosca and Pareto emphasized elites' need to replenish their ranks through constant but gradual and peaceful inflows and outflows of members. Pareto also toyed with a more dramatic process, in which cunning and foxlike elites are replaced, gradually or suddenly, by belligerent and lion-like elites, who are, however, in turn eventually displaced by new foxlike elites in an endless cycle. In this context, Pareto famously depicted revolutions as amounting to little more than circulations of ruling elites. Processes of elite circulation and transformation appear to vary according to the basic type of elite that exists in a society. Where this is a disunited elite, sudden and violent displacements of ruling elite camps and cliques tend to occur frequently. Through a coup or elite-led popular uprising, one camp or clique seizes and holds power until the next coup or uprising. Modern history is replete with such violent but essentially garden-variety elite circulations. More rarely, the ascendant camp in a disunited elite may so mismanage its rule or suffer so many misfortunes that the state and its coercive forces collapse, opening the door to a revolutionary interregnum and the victory of a previously peripheral but tightly organized and doctrinaire elite group. Where this happens, one may speak of a transformation from the disunited to the ideologically united type. The classic example is the collapse of the war-ravaged Tsarist state and the ensuing revolutionary triumph of the tiny Bolshevik elite in Russia between 1917 and 1921.

Where, as in the Soviet Union from 1922, an ideologically united elite holds sway, elite circulation occurs more or less exclusively by jockeying for high positions in a crushingly dominant party or movement. Eventually, however, the imposed and "official" ideology or other dogma of that party or movement may become so threadbare and at odds with societal developments that its elite unravels or implodes. The ideologically united elite is transformed into a disunited elite, different in makeup but not in basic type from the disunited elite that preceded the earlier revolutionary transformation. This was the fate of the Soviet Union's ruling elite some 70 years after it took power. Its implosion created disunited elites in Russia and all of the newly independent states that had constituted the Soviet Union. A question at present is whether similar implosions of ideologically united elites will occur in China, Cuba, Iran, North Korea, Vietnam, and perhaps Iraq.

Consensually united elites appear to be immune to such dramatic circulations and transformations. Sudden and irregular seizures of governing power by competing elite camps and factions do not occur, and the historical record indicates that, once formed, a consensually united elite strongly tends to persist over long periods: in England/Britain since 1689, the United States since 1789, Sweden since 1809, Canada since 1867, Australia since 1901, India since 1947, inter alia. Elite circulation occurs gradually through periodic election contests and complex career and other hierarchical recruitment structures in which "merit" is emphasized and rewarded, even if mainly in the eyes of those who are successful. This circulation is not inconsequential, however. Modern communication technologies so heighten elite organization and mobilization of pubic support that electoral competition and policy disputes increasingly approximate battles between armada-like coalitions of elites, whose alternating victories bring sharp changes in governing style and policy content.

Elites and Political Science

The study of elites confronts political science with difficult problems. First, elite behavior always contains enough arbitrary and capricious elements to defy deterministic explanation. A theory of elites and politics must allow considerable leeway for unpredictable elite choice, and its explicative claims must, therefore, be modest. This means, second, that a theory should focus principally on what is unlikely or impossible in politics—on what the inescapability of elites and their relatively autonomous actions rule out. Yet much political science regards political possibilities—especially possibilities for making people more free and equal, more enlightened and participative—as essentially open-ended.

Third, the secrecy that elites routinely practice is at odds with the strong empirical bent of political science. Compiling more data on elite backgrounds and other aspects of their profiles, conducting more surveys of usually evasive elite opinions and momentary network locations, and manipulating such data with rigorous statistical techniques are unlikely to bring political scientists appreciably closer to a robust understanding of elites and politics. In short, elites often do not lend themselves to empirical and inductive research. Capturing what elites do and why they do it requires a deductive and rationalistic theory that must be tested against observed events. Do events occur as they would if the theory were true? Does the theory make better sense of what happens politically than other theories? If it does, are there nevertheless clear and undoubted facts that show the theory to be at least in part mistaken? These are legitimate and necessary questions. "Can each of the theory's postulates about elite behavior and motivation be demonstrated and measured empirically?" is not, however, a legitimate question. No seriously explicative theory, even in the most experimental and reputedly scientific fields, can meet such a test. Things that are observable merely happen as they would if the theory were true and no more elegant and parsimonious explanation is available. While barriers to empirical research and inductive explanation exist in many areas of political science, they are especially pronounced in the study of elites.

John Higley
University of Texas at Austin
Austin, Texas, United States

See also Democratization; Elitism; Interviews, Elite; Masses; Oligarchy; Political Class; State Formation

Further Readings

Albertoni, E. A. (1987). *Mosca and the theory of elitism*. London: Blackwell.

Best, H., & Higley, J. (Eds.). (2009). *Democratic elitism in transition*. Leiden, Netherlands: Brill.

Cotta, M., & Best, H. (Eds.). (2007). *Democratic representation in Europe: Diversity, change, and convergence*. Oxford, UK: Oxford University Press.

Dogan, M. (Ed.). (2003). *Elite configurations at the apex of power*. Leiden, Netherlands: Brill.

Femia, J. (2007). *Pareto and political theory*. London: Routledge.

Higley, J., & Burton, M. (2006). *Elite foundations of liberal democracy*. Lanham, MD: Rowman & Littlefield.

Linz, J. J. (2006). *Robert Michels, political sociology, and the future of democracy*. New Brunswick, NJ: Transaction Publishers.

Parry, G. (2005). *Political elites* (2nd ed. with a new introduction by the author). Essex, UK: ECPR Classics. (Original work published 1969)

Putnam, R. D. (1976). *The comparative study of political elites*. Englewood Cliffs, NJ: Prentice Hall.

Schumpeter, J. (1942). *Capitalism, socialism and democracy*. New York: Harper.

Elitism

The etymology of the word *elite* (occasionally spelled *élite*) comes from the Latin *eligere* (to elect). In general usage, an elite is a relatively small dominant group or social category of people who occupy a position of privilege or dominance and have a privileged status within a larger society. To examine the "elite" category in political science is to take up a fundamental question—that of defining the group or groups that hold "power" in a social system, that is, those who have a dominating influence on the definition and the carrying out of public decisions. The question of who holds power in modern societies is generally dealt with from three different theoretical standpoints: (1) elitism, (2) pluralism, and (3) Marxism. From an elitist perspective, power belongs to an elite that is rather united. Pluralists, on the other hand, consider that there is no single ruling class but rather a plurality of ruling groups that alternate between cooperation and confrontation. According to Marxists, power in capitalist societies is monopolized by a dominant class—the bourgeoisie—and its auxiliaries. Here, domination is concealed by ideology.

Theories of Elitism

The concept of the elite was first introduced by Vilfredo Pareto (1848–1923), an Italian economist, sociologist, and senator. A firm believer in economic liberalism and an adversary of socialism, he aimed at challenging the Marxist conception of class struggle. In 1901, he published *Rivista italiana di sociologica*, a famous text later translated as *The Rise and Fall of the Elites: An Application of Theoretical Sociology* (1968), and in 1916 he published *Trattato di Sociologia Generale*, later translated as *The Mind and Society* (1935). He defined members of the elite as people with exceptional virtues who show distinguished abilities in any domain. Being part of the elite therefore depends on individual capacities and natural talent that lead to above-average success.

Gaetano Mosca (1858–1941) had previously worked on a very similar theory, although not explicitly using the concept of *elite*. A professor of public law, and deputy and then life senator of the former Kingdom of Italy, he wrote the first version of *Elementi di scienza politica* in 1896 (*The Ruling Class*, 1939). He further developed the notion of the "ruling class" and the idea that all societies are based on a distinction between the ruled and rulers. According to Mosca, the former is always an organized minority that takes on all political responsibilities, monopolizes power, and enjoys all the subsequent advantages. The latter, always greater in number, is led and controlled by the rulers, who act more or less legally and sometimes violently.

The German socialist Roberto Michels (1876–1936) is also among the elitist authors. A pupil of Max Weber, he studied in England, France, Germany, and Italy, where he taught economics and political science. After World War I, he joined Mussolini's Fascist party, whose ideals he perceived as a more democratic form of socialism. In 1911, he published *Zur Soziologie des Parteiwesens in der modernen Demokratie. Untersichungen über die oligarchischen Tendenzen des Gruppenlebens* (*Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracies*, 1915). His premise was that the masses lacked the "mechanical and technical" capacity to govern themselves. According to Michels, a society cannot exist without a dominant class because leadership is technically indispensable to the survival of organizations. The principle of his "iron law of oligarchy" is that one dominant class inevitably succeeds to leadership in an organization, as the collective psychological characteristics of both masses and elites make the leaders of any organization become concerned above all with remaining in office and strengthening their control over the organization.

In 1941, when his seminal work *The Managerial Revolution: What Is Happening in the World* was published, James Burnham (1905–1987) had already left the American Socialist Workers Party and was soon to serve as a public intellectual of the American conservative movement. Burnham argued that the 19th-century bourgeoisie has been replaced since World War I by a "social group or class," which he called "managers." They control the management, the leadership, and the coordination of the economy. But these managers are not capitalists as they do not own, finance, or have a major share in the company. Thus, the main difference between the ruling elite and the masses does not lie in ownership as much as in the control of the means of production, in particular through knowledge and technical competence. According to Burnham, managers do not seek profits so much as they seek power itself, their incentive being more political than economic. They do not distinguish between corporate and state ownership, and their oligarchy holds power in both capitalist and socialist regimes.

These four authors are sometimes called the "Machiavellians," as they are deemed to convey a pessimistic and cynical view of the world. They defend the idea of the autonomy of politics toward economic infrastructure and give a prominent position to political factors (and not economic ones), as their main interest lies in power, not in ownership.

Usually defined as one of the elitist authors, the American sociologist C. Wright Mills (1916–1962) opposed both Marxism (and its principle that power is based on the ownership of the means of production) and the liberal version defended by the previous authors (in particular, the idea of an autonomous political order). Mills (1956) showed that in the United States, the main decisions are made by an elite composed of the people at the top of the three big institutional hierarchies (political, military, and economic elites) that form a "triangle of power." These major institutional orders are interdependent and interlocked, and their members share a common worldview and a class identity (they consider themselves as being separate and superior to the rest of the society).

Similarly, the American professor of psychology and sociology G. William Domhoff (1936–) argued in his 1967 book *Who Rules America?* that there exists a ruling class in the United States that is the elite and in it is concentrated the major part of national wealth; it controls the major companies and the main banks; runs the foundations, universities, and media; and therefore leads the country's economy. According to Domhoff, those who own income-producing property (corporations, real estate, and agribusinesses) set the rules within which policy battles are waged. For him, economy is the only power network of any consequence in the history of the United States. Floyd Hunter (1953) reached similar conclusions about power at the local level. He examined in detail the power relationships in Atlanta, Georgia, where he found that the city was led by a small and homogeneous elite and that "real" decisions were made by a corporate oligarchy.

Elites' Circulation

The notion of elites' circulation is central to elitist theories. According to Pareto, because being part of the elite is not hereditary, there is a constant replacement of old elites by new ones via the circulation of individuals, a process that enables a certain social equilibrium (and prevents revolution). This continuous replacement of one elite by another is the driving force behind history. Mosca further developed Pareto's theory of elites' circulation, claiming that modern democratic societies enabled a high degree of upward social mobility. For these authors, hereditary inequalities are secondary and affect elites' circulation only marginally. Individual capacities are indeed necessary to remain part of the elite, given the fierce individual competition at stake.

According to Mills, the interlocking nature of the power elite allows for interchange between different sorts of institutions, maintaining and strengthening the power of each. But circulation is confined to the "three bigs," among which it is easy to move frequently given the shared interests and origins of the rulers. In the same vein, the French political sociologist Pierre Birnbaum (1940–) holds that there is interpenetration and a never-ending circulation within this socially and culturally homogeneous ruling space (Birnbaum, 1978).

Criticisms of Theories on Elitism

The Pluralist View

Strongly inspired by Weber's work, supporters of pluralist analysis believe that public policy is the result of interactions between the players and various competing groups situated both inside and outside the sphere of the state. No particular group, or elite, can occupy a constant and far-reaching dominant position.

Robert Alan Dahl (1915–) strongly opposes Mills's and Domhoff's views on the nature of politics in the United States. Dahl holds that far from being unitary and demographically narrow, the elites are numerous and autonomous, and alternate between competition and compromises. According to his polyarchy model, no elite dominates, which makes compromises and negotiations necessary. His 1961 study of the formal and informal power structures in New Haven, Connecticut, where the University of Yale is located, supports this view. This analysis can also be found in Aaron Wildavsky (1930–1993) who examined local decisions made in Oberlin, Ohio. In his 1964 book, *Leadership in a Small Town*, he concludes that the structure of power is pluralistic and that according to the nature of the problems at stake, different groups will lead and rule at the local level.

According to the French sociologist, political scientist, philosopher, and journalist Raymond Aron (1905–1983), power is structured in a very complex way in modern societies (Aron, 1950). He distinguishes the political elite from five ruling categories: those who hold "spiritual power" and influence the ways of thinking and believing (priests, intellectuals, scientists and scholars, and ideologists), army and police chiefs, managers of collective work (those who own or control the means of production), mass leaders (at the head of trade unions or political parties), and senior civil servants who hold administrative power. In modern societies, these categories are highly competitive.

The Marxist Thesis

Marxist theory is based on an economic analysis according to which power is held by the owners of the means of production, the bourgeoisie, who also monopolize political power and control the state apparatus. According to Marxists, economic factors play a prominent role; the state, far from being neutral, is a means of political domination in the hands of the bourgeoisie.

According to Marxists, pluralists conceal reality by substantiating the liberal illusion of an autonomous political order. They give credence to the idea of state neutrality and mask the true nature of state power as well as class struggle. Marxists also criticize Mills's elitism for not being based on the ownership of the means of production. Elitists in general are criticized for focusing on political domination and failing to identify it with the economic class that owns the means of production. This criticism is in particular supported by the Greco-French political sociologist Nicos Poulantzas (1937–1979), who rejected the very notion of elite in *Pouvoir politique et classes sociales* (1968), published in English as *Political Power and Social Classes* (1973). He disagreed with another Marxist political theorist and sociologist, Ralph Miliband (1924–1994), a friend of Mills, who claimed that the notion of elite was relevant and that Marxists could even admit the plurality of elites. Besides being diverse, these elites all belong to the same dominant social class and share the same social origins, the same networks, and the same ideology.

Elites and Socialization

For a theorist such as Mills, leaders share the same origins, the same interests, and the same education and training. Their solidarity stems from their "social similarities and psychological affinities," in which their common training plays a critical role as the basis for permanent alliances of power and for strategic marriages and other affiliations. According to Domhoff (1967), educational institutions participate in the reproduction of the power of the wealthy few. Schools and universities are indeed developed by the upper class and are filled with their children, who are socialized in an upper-class worldview along with the newly wealthy people who are assimilated there. The primarily White and Protestant elite of the Ivy League that he identifies as being dominant in all fields hold a common vision of society and "what must be done" to govern. This homogeneity is linked to a homologous socialization (same schools, same social circles, etc.) and to their easy access to positions of power both in and out of the government.

Education is also critical in the reproduction of the social order, according to the French sociologist Pierre Bourdieu (1930–2002). He showed that elites are able to reproduce themselves, that is, pass on their privileges to their offspring, even in the ostensibly most meritocratic social fields, such as education. The educational system enables children from the dominant class to obtain the best diplomas and thus gain access to dominant social positions. At the same time, the system legitimizes their academic achievement by linking it to individual and natural gifts, concealing its social origins (Bourdieu, 1996).

Education and Monopolization of Power Positions

According to Birnbaum (in Madeleine Grawitz & Jean Leca, 1985), the differentiation between state and society, which is greater or smaller depending on the country and the time period, may account for both the contours of the elite or elites involved in the making of public decisions and the extent of their power. Thus, for example, in the case of France, at least since 1945, the combination of a protean state intervention and networks for the education and selection of elites in a few prestigious state "*grandes écoles*" have contributed to the quasi-monopolization of ruling positions by those belonging to the highest levels of the engineering and administrative fields (Ezra Suleiman, 1978). This tendency was mainly reinforced by the relative weakness of intermediary bodies, employers' organizations, and trade unions and, from the 1960s onward, by a process of political "technocratization," that is, the access to positions of political responsibility by high-ranking civil servants. The fact that these situations were not called into question undoubtedly contributed to the legitimization of an ultimately hegemonic organization of power distribution.

If one believes that the place occupied by the elite(s) varies according to the problems presented and the time period concerned, then the beginning of the crisis of the French model of public policies in the 1970s had a significant effect on the elitist structures. Members of the state nobility still occupy the majority of the dominant positions in the field of public service and continue to be the main players. The ability of the administrative elite to move easily from one social group to another (e.g., to hold directorial positions in large private firms) allows them to continue to capture a lot of power, even if they are not *in* power. But a state's fall into a crisis, on both ideological and practical levels, is also accompanied by the promotion of new elites who are supposed to embody new ways of defining public policy. The more competitive nature of the decision-making networks, the introduction of private sector–inspired management models into public administration, Europeanization, and even the globalization of issues, and thus of procedures, are directly contrary to a unified, monistic conception of power. Administrative elites are thus beginning to be seriously challenged by corporate actors such as private consultants, who have been educated in the same "*grandes écoles*" and impose themselves as experts on public sector matters, as in other countries (Canada, the United States, or the United Kingdom) where their success has been much quicker (Denis Saint-Martin, 2001).

The Concept of Elite and the Issue of Democracy

Most authors agree on the fact that democracy is an illusion and impossible to achieve. According to Roberto Michels, history is led by elites' struggle for power. His idea of the iron law of oligarchy was the result of his analysis of leadership patterns in democratic governments and in other organizations, such as trade unions. According to this law, the overthrow of any elite by the masses will eventually lead to such oligarchical leadership. Michels, therefore, refutes the possibility of the existence of an ideal democracy. For the supporters of classical elitist theories, the existence of elites at all times and places is a natural phenomenon because of unequal distribution of talent, wealth, and political influence in society.

According to Mills, the great masses of people are largely unorganized, ill informed, virtually powerless, and controlled and manipulated by the powerful elite who exploit them both economically and politically. Their dependence and disorganization prevent the masses from participating in democratic life. Above the masses, Mills saw a middle level of power (composed of local opinion leaders and special interest groups) that is reflected in the U.S. Congress and in U.S. political parties. But they not only fail to represent the masses but also have no real effect on the elites, who leave them to debate and decide some minor issues, as long as they do not represent a serious challenge to their authority. The political directorate itself is described as undemocratic in both the process of its selection and its maintenance. Moreover, Mills believed that widespread alienation of the masses, their political indifference, and the economic and political concentration of power are serious threats to democracy. The U.S. political scientist Harold Dwight Lasswell (1902–1978), who also used both power and control of hierarchies as elements in identifying elites, insisted on the critical importance of legitimacy. In his 1930 book *Psychopathology and Politics*, he held that the extent of elites' positions of authority is determined by the reactions of the public toward those who hold power. The existence of political elites depends on the legitimacy of their authority "in the eyes of the people."

Pluralists do not believe in the ideal of democracy either. For Dahl (1989), it is a theoretical utopia. But Dahl sees polyarchical institutions as a major improvement as they create multiple and autonomous centers of political powers where decisions can be made through competition, bargaining, and compromises.

Marxists also view liberal democracy as an unrealistic utopia. According to them, the capitalist state is inherently undemocratic as it represents the dictatorship of the bourgeoisie. The fact that intermediary powers such as the media and most political parties need the support of the bourgeoisie to win elections makes the system intrinsically undemocratic. Vladimir Lenin asserted in *The State and Revolution* (1917) that the dictatorship of the proletariat would be the highest possible form of democracy.

Broadening and Completing the Analysis

In other respects, the development of research on the structuring of public decisions and their practical implementation has led to a better understanding of the role played by the elite. The classic idea that it is the position held that infers the power now appears to be a necessary but insufficient condition. Research on the sociology of elites conducted in the 1960s and 1970s remains quite relevant when it comes to describing the social and educational resources required to hold a position of influence and/or power. On the other hand, since the 1980s, developments in sociology of public policy have allowed for the reexamination of at least three avenues of research:

1. Public decisions are made in an often competitive and uncertain context. Holding a position objectively seen as being dominant does not guarantee the ability to effectively develop a given public policy. If the "dominators" remain dominant, they must exercise their authority more often than in the past and bargain to keep it.
2. Exerting a direct influence on the definition of a public policy does not necessarily mean that the policy, once implemented, will not be the object of various and sometimes unexpected reappropriations that can ultimately produce concrete results that are quite different from what the producers intended.
3. To understand public policies through the eyes of an elite(s), one must go beyond the question of specific interests to take into account the values, world visions, and, more generally, the knowledge that together play a role in the framing of political exchanges and decision making.

One of the major contributions of the sociology of the players and the studies of "street-level bureaucracy" (Michael Lipsky, 1980) or middle-ranking officials (Herbert Kauffman, 1960; Page, 2003) who carry out the wishes of the administrative elite is to show that public policies are shaped and designed during the implementation process by people with low- or middle-range powers. In short, the elitist approach to public policy, since it expresses a global point of view, can be supplemented and enriched by a more microsociological perspective that provides a better explanation of public policies in action and thus proposes a re-reading of elitist theories in the negative.

Pierre Mathiot
Institut d'Etudes Politiques de Lille
Lille, France

Julie Gervais
Triangle Centre for Political Research, CNRS
Lyon, France

See also Class, Social; Elites; Oligarchy; Political Socialization; Power

Further Readings

Aron, R. (1950). Social structure and the ruling class. *British Journal of Sociology*, *1*(1–2), 126–143.
Bourdieu, P. (1996). *The state nobility: Elite schools in the field of power* [*La Noblesse d'Etat: Grandes écoles et esprit de corps*, 1989]. Stanford, CA: Stanford University Press.
Burnham, J. (1941). *The managerial revolution: What is happening in the world.* New York: John Day.
Dahl, R. A. (1961). *Who governs? Democracy and power in an American city.* New Haven, CT: Yale University Press.
Dahl, R. A. (1989), *Democracy and its critics.* New Haven, CT: Yale University Press.
Domhoff, W. (1967). *Who rules America?* Englewood Cliffs, NJ: Prentice Hall.
Kaufman, H. (2006). *The forest ranger: A study in administrative behavior*. Washington, DC: Resources for the Future. (Original work published 1960)
Lasswell, H. (1930). *Psychopathology and politics.* Chicago: University of Chicago Press.
Lipsky, M. (1980). *Street-level bureaucracy: The dilemmas of individuals in public services.* New York: Russell Sage Foundation.

Michels, R. (1915). *Political parties: A sociological study of the oligarchical tendencies of modern democracies* (E. Paul & C. Paul, Trans., from *Zur Soziologie des Parteiwesens in der modernen Demokratie. Untersichungen über die oligarchischen Tendenzen des Gruppenlebens*, 1911). New York: Hearst International.

Mills, C. W. (1956). *The power elite.* Oxford, UK: Oxford University Press.

Mosca, G. (1939). *The ruling class* [*Elementi di Scienza Politica*, 1896]. New York: McGraw-Hill.

Page, E. C. (2003). The civil servant as legislator: Law making in British administration. *Public Administration, 81*(4), 651–679.

Pareto, V. (1935). *The mind and society*. New York: Harcourt Brace.

Pareto, V. (1968). *The rise and fall of the elites: An application of theoretical sociology* ["Un applicazione die teorie sociologiche," *Rivista Italiana dieSociologia*, 1901, 402–456]. Totowa, NJ: Bedminster Press.

Poulantzas, N. (1973). *Political power and social classes* [*Pouvoir politique et classes sociales*, 1968]. London: New Left Books.

Saint-Martin, D. (2001). *Building the new managerialist state: Consultants and the politics of public sector reform in comparative perspective.* New York: Oxford University Press.

Scott, J. (1991). *Who rules Britain?* London: Polity Press.

Suleiman, E. (1978). *Elites in French society*: *The politics of survival.* Princeton, NJ: Princeton University Press.

Wildavsky, A. (1964). *Leadership in a small town.* Totowa, NJ: Bedminster Press.

EMPIRE

An empire is a type of political organization in which the metropolis exercises control over diverse peripheral actors through formal annexations and/or various forms of informal domination. This short definition may well be uncontroversial, but when one tries to make it more specific consensus is lost. This is partly because most social science categorizations reflect abstract models that appear in a variety of shapes, forms, times, and places. In this sense, the term *empire* is as fuzzy and contestable as terms such as *democracy*, *class*, *power*, or *nation*. But unlike these other categories, *empire* is laden with negative historical and ideological connotations and is often used as a synonym for colonial rule based on oppression and exploitation. This entry examines the concept of empire, the historical evolution of empires, and the ways in which a study of empire can shed light on international politics.

The Nature and Evolution of Empire

Such an understanding of empire is ahistorical and biased. Empire has been a characteristic form of political organization since early antiquity and predates colonial rule by several centuries. The colonial era may be over, but this does not imply the end of empire (although it is important to emphasize that the colonial legacy still haunts former colonial empires and their erstwhile colonies). Moreover, many empires were symbols of peace and prosperity, rather than of oppression and exploitation. Studies of empires show that control within them can be based on incentives rather than coercion or on a combination of both. It can be exercised through a variety of military, economic, and cultural means. It can be formal or informal to varying degrees. Periphery status within an empire can also differ. Some peripheral actors are given access to the decision making and resources of the metropolis, while others are kept at a distance or even subject to open discrimination and exploitation. The relationship between metropolis and periphery can certainly be hierarchical and conflict ridden, but it can also be harmonious and based on mutual dependency. Some empires are organized in concentric circles or even form quite loose multiple independencies.

The nature of both metropolis and peripheral actors can also differ. In most cases, the metropolis has a centralized government, differentiated economy, and shared political loyalties, while the "imperializable peripheries" (to use Michael Doyle's term) have weak government, undifferentiated economies, and highly divided political loyalties (Doyle, 1986, p. 19). However, the imperial metropolis can also have a relatively weak, limited, and decentralized government; an inefficient economic system; and multiple cultural identities. For instance, medieval empires are known for having limited and decentralized government performing only a few basic governmental functions. They were ridden by internal conflicts between king (or emperor) and the lower aristocracy (whether feudal or bureaucratic), while the persistent divergence of local cultures, religions, and traditions implied a highly divided political loyalty.

The metropolis does not always have a master plan of imperial conquest. States can become empires by default because they try to bring some order to unstable neighbors or try to convert barbarians into "good" citizens or into a specific religion. Likewise, an empire does not necessarily come into being through outright aggression. Some empires emerge quietly or even surreptitiously through uneven modernization and social differentiation. They may not even see themselves as empires.

Different historical periods have generated different types of empires. Alexander the Great built his empire and brought a multitude of nationalities under his central authority largely by the use of military force. The Roman Empire relied far more on peaceful methods, using language and law in the service of its civilizing effort. The Chinese Han Empire was renowned for its administrative skills, enabling it to keep diverse and autonomous provinces under a single rule. The British Empire, like those of France, Portugal, and Spain, used its maritime power and supremacy in global trade. The Soviet Empire skillfully applied ideological penetration alongside the older military, political, and economic techniques of empire building.

The "end of empire" has been declared several times throughout history, but at the beginning of the 21st century it looks as if empires are on the rise rather than in decline. The United States, China, the European Union (EU), and Russia are most frequently described as empires. All four represent vast territorial units with global influence in material, institutional, and ideological terms. They possess not only global economic and military reach but also the ability to influence the global institutional agenda and shape the notion of legitimacy (if not normality) in various parts of the world. In addition, all four have a record of acting in a way that imposes significant domestic constraints on a variety of formally sovereign actors. These actors are seen as a kind of periphery to be governed by the imperial center. (China and Russia also contain peripheries within their own borders and without formal sovereignty, such as Tibet or Chechnya.) Moreover, the policies of the United States, China, the EU, and Russia are guided not merely by their selfish interests but also, if not primarily, by a universal normative vocation reminiscent of the *Pax Romana* or the old Chinese *Mandate of Heaven.* In other words, they all pursue a civilizing mission to benefit their region if not the entire world. John Ikenberry notes that the United States claims to be an "indispensable nation" maintaining global order and possessing an economic project "congruent with the deeper forces of modernization." The EU sees itself as a pole of attraction for its neighbors, contributing to what the EU High Representative has described as a fairer, safer, and more united world. Even contemporary Russia sees itself as a bastion against the barbarian forces of chaos, nationalism, and religious fundamentalism flourishing in its backyard.

However, there are significant, and in some cases striking, differences among these four contemporary empires. Consider, for instance, the difference between an autocratic, nationalistic, and militaristic Russia and the peculiar, postmodern polity that is the EU, with no single center of government, no single *demos* (let alone nation), and no army. Nor is there an American equivalent of a subjected inner periphery like Tibet, Xinjiang, and what China calls Inner Mongolia. However, many of those who study empire argue that the United States tends to treat the entire world as its periphery. Its scope of territorial control is truly global. Although control by the United States is usually informal, it is backed up by enormous military and economic might.

Of course, empires may rise and fall even while the criteria for defining them are under discussion. Japan, India, Brazil, South Africa, and Germany may be seen as empires of a sort or may become empires in the next few decades. For instance, Japan's reluctance to reassume an imperial posture is more a function of historical memory than of military capability. Contemporary Japan, like China, shows signs of a "middle kingdom" mentality—a belief that the world revolves around it. This is clearly an imperial feature. And Japanese military expenditure is nearly as large as the Chinese and much higher than the Russian.

India also shows signs of a middle-kingdom mentality. Moreover, India exercises significant influence over Sri Lanka, Nepal, and Bangladesh. Time and again, India has been able to use its power quite effectively in global trade talks, environmental negotiations, and interstate disputes. Many Indians believe that their democratic and economic profile, together with their language abilities, give them the edge over Russia or China. But the strength of Indian society has yet to translate into strong institutions, and this hampers India's global rise and contains its imperial aspirations.

Empire and the Analysis of International Politics

If the list of empires is open-ended and the differences between them are so striking, does a study of empire help our understanding of international politics? Is the state paradigm not sufficient and, indeed, more suitable? The problem is that the notion of a state is as fuzzy as the notion of empire. The United States is a state; but so too are Somalia, Luxembourg, and Costa Rica. Moreover, the state paradigm, with its emphasis on territoriality and sovereignty, seems ill suited for understanding the 21st century with its cascading economic interdependence, transnational jurisdictions, and multiple cultural loyalties. Besides, the reemergence of empires does not herald the end of nation states. As Herfried Münkler (2007) pointed out, contemporary imperial structures are superimposed on the state order, but they no longer replace it: "Whoever thinks of imperiality as simply an alternative to statehood will come to the conclusion that no empires exist today" (p. 6).

Some authors prefer to use the term *hegemony* in place of *empire*. However, hegemony is not necessarily a synonym for empire. According to Alejandro Colás (2007, p. 165), for example, hegemony, unlike empire, has no final instance, no decisive location or moment of authority, but is reproduced instead through a sequence of coordinated and generally consensual actions by a multiplicity of actors, albeit orchestrated by a leading state. True, the concept of hegemony, like the concept of empire, underlines the inequality among individual actors and the ensuing hierarchy in the international system. However, the literature on hegemony tells us less about the nature of the actors exercising hegemony than does the literature on empire. Power is often detached from agency in the study of hegemony. Hegemony discourse is more about the symptoms of power politics than its causes.

The study of empire is not only about power but also about the actors engaged in power politics. As such, it helps us understand the nature of the magnetic poles within the current global system. Moreover, it forces us to examine power not only in crude economic or military terms but also in institutional and ideological ones. As noted above, most empires try to "civilize" and "institutionalize" their peripheries, not merely to conquer them. The study of empire also underlines the role of order, hierarchy, and structure in international politics. It is not chiefly about exploitation and inequality, as is often the case in the study of hegemony. In fact, the study of empire shows that some peripheries can display an impressive ability to determine the politics of the center, despite all the material or normative asymmetries. This leads to another advantage that a focus on empire offers: It helps overcome the often misleading distinction between domestic and foreign affairs. Empires seldom possess fixed and hard borders, and this fuzziness or permeability of borders largely determines the politics of the imperial center. For instance, the nature of European integration is constantly being reshaped by its successive waves of enlargement. Russian policy is largely determined by the fate of ethnic Russians living outside its formal frontiers. The one-way permeability of U.S. borders allows Washington to intervene in the politics of various states in Central America and the Caribbean, without having seriously to consider the possibility that they will intervene in the U.S. sphere.

Finally, the study of empire helps us understand the evolving pattern of cooperation and conflict among major power poles. History shows that empires can coexist in parallel if they do not contest similar spheres and spaces and if they respect each other's claims to legitimacy. The parallel functioning of the Roman and Chinese empires is a good example of such coexistence. Contemporary empires often contest similar spheres and spaces. Moreover, the dominance of the United States is likely to generate conflict even though other contemporary empires have not yet tried to balance or deter the United States.

Jan Zielonka
University of Oxford
Oxford, United Kingdom

See also Colonialism; Hegemony; International System; Power; State

Further Readings

Colás, A. (2007). *Empire.* Cambridge, UK: Polity.

Doyle, M. W. (1986). *Empires.* Ithaca, NY: Cornell University Press.

Maier, C. S. (2006). *Among empires: American ascendancy and its predecessors.* Cambridge, MA: Harvard University Press.

Motyl, A. J. (2001). *Imperial ends: The decay, collapse, and revival of empires.* New York: Columbia University Press.

Münkler, H. (2007). *Empires: The logic of world domination from ancient Rome to the United States.* Cambridge, UK: Polity.

Zielonka, J. (2006). *Europe as empire.* Oxford, UK: Oxford University Press.

EMPOWERMENT

Empowerment is a concept that has received much interest in the generic management literature during recent years. There are numerous definitions of this concept, but the common denominator of all these deals with the distribution of power to lower levels of a hierarchy. The meaning of empowerment for the public sector is, however, an arena that calls for more elaboration and revivification. This short entry aims at three major goals: (1) to provide an up-to-date review and explanation of the generic meaning of empowerment, (2) to explore the importance of empowerment for public sector domains, and (3) to provide a short direction to different foci of empowerment in the context of the public sector.

The Essence of Empowerment: A Broad Organizational and Managerial View

Empowerment is a multifaceted construct, and there is much deliberation with regard to its meaning and focus. Discussions regarding its meaning include questions of whether it is merely the sharing of power, a more general concept of decentralization, or a psychological motivational concept regarding the person's feelings of worth. The focus or target of the empowerment process may be employees, the end users/clients/public, or institutions/organizations/subnational governments.

As control and power are regarded in the management literature from either a relational perspective or a motivational perspective, the concept of empowerment too can be seen from these two different points of view. From the relational perspective, power is a function of dependence and/or interdependence of parties, and thus empowerment is a process in which a leader or manager shares his or her power with subordinates, making them less dependent on him or her. In accordance, empowerment is practiced, by those adopting this perspective, through participative management techniques such as goal setting by subordinates and quality circles. From a motivational perspective, power is an intrinsic need for self-determination or a belief in personal self-efficacy. Thus, empowerment is the creation of conditions enhancing a strong sense of self-efficacy. Kenneth Thomas and Betty Velthouse (1990) built on this psychological-motivational definition and argued that empowerment is a multifaceted concept of four cognitions: (1) meaning, (2) competence, (3) self-determination, and (4) impact. Meaning is a match between the job requirements and the person's beliefs, values, and behaviors. Competence is the person's belief in his or her own capability to perform the job well. Self-determination is related to autonomy in initiating and continuing work behaviors and processes. Finally, impact is the extent to which a person can influence strategic, administrative, or operational outcomes. These cognitions are shaped by the person's work context and personality characteristics and in turn motivate individual behavior. Thus, empowerment processes under this perspective will encourage self-managing or cross-functional teams and will provide information about the organization's goals and strategy to give the workers a sense of ownership. Similarly, according to David Bowen and Edward Lawler (1995), empowerment entails a high-involvement approach. Thus, organizations are encouraged to change their policies, practices, and structure in order to create and sustain empowerment, and this change must distribute power, authority, responsibility, information, knowledge, and rewards throughout the organization.

The Meaning of Empowerment in the Public Sector

The generic psychological understanding of empowerment in organizations is useful and meaningful when state agencies and social governmental environments are discussed. Essentially, the discussion of empowerment in such neighborhoods calls for integration of other concepts such as democracy, citizens' involvement, public sector reforms, and multiple stakeholders' engagement. As will be explained below, such a discussion must also consider both the personal and the institutional implications of empowerment. Putting it another way, what is the contribution of empowerment to the theory and practice of public administration, to public organizations, and to the overall understanding of government and governance?

Most important, empowerment can be seen as a democratic strategy that involves multiple stakeholders in the action of governance or in the action of managing governmental agencies. As such, empowerment allows greater self-contributing behavior; it encourages individuals' engagement in public actions and advances the production of public goods and services. It can also contribute to commitment to public service and to public sector motivation (PSM) by all those who are part of the process, be it public personnel, citizens, residents of cities and communities, or other stakeholders such as the private sector and the third sector.

There are two major perspectives of empowerment in this regard. One perspective sees democracy as a predominantly aggregative theory, and the other perspective sees democracy as a predominantly integrative theory. Predominantly aggregative theories see democracy as a means of distributing political power and influence and of regulating conflict within a society. Thus, empowerment refers to equal influence over processes of collective governance and individual autonomy. According to Albert Hirschman, citizens have two means of empowerment available to them when they are not satisfied with the consequences of a collective action: exit and voice. In an aggregative strategy of empowerment, exit by voting for a new political party is a way of punishing politicians and political organizations that were unsatisfactory and a way of escaping unacceptable majority decisions. Voice, in this perspective, is a supplementary means of empowerment as it gives the decision makers information about the reasons exit was taken.

Predominantly integrative theories see democracy as a means for producing democratic citizens. According to this view, society is a social construction that is more than the sum of the preferences of individuals. Empowerment is therefore the transformation of individuals into citizens by participation. Thus, in this perspective, voice is the main vein of empowering individuals as it gives them a chance to participate in decision-making processes. As Danny Burns, Robin Hambleton, and Paul Hoggett (1994) point out, voice empowers as it informs, organizes, and gives confidence to the citizens. They argue that a strategy of empowerment under this perspective must include decentralization of authority to local governments, institutionalization of communication channels between citizens and politicians, and encouragement of citizen participation in general.

The Target of Empowerment: A Triple Meaning

Based on the above, empowerment in public administration employs at least three major targets: (1) public officers/employees, (2) citizens as customers, and (3) institutions or subsectors. Whereas the first relate to the individual level, the latter refers to the institutional and subgovernment level.

First, stemming from the generic managerial literature, the most classical orientation of empowerment in the public sector is with public sector employees. These internal "agents" of policy and public action are the primary mechanism through which public strategies, policies, and actions are implemented. Public employees and public officers can be trained and encouraged to take greater control over their jobs, tasks, and decisions in the organizations. This process of empowering people should go hand in hand with formal rules and procedures that characterize the public bureaucracy. Although they cannot go against them, they must also give employees more flexibility in doing the job. Flexibility does not mean taking individual decisions that contradict the formal rules that bureaucracies so typically maintain and safeguard. Hence, although bureaucracy may become some barrier to empowerment, there is much room for building innovative mechanisms that allow healthy involvement, intra-organizational democracy, and self-motivation that works for the good of many. The restraints and control of bureaucracy also somewhat limit the voice of employees, but modern managerial techniques such as team building, self-managed teams, job enrichment, and even quality circles or goal-setting strategies may increase commitment and build a sense of recreation in public sector agencies.

Next is the citizenry level. While in the management literature the focus or target of empowerment are the employees, in the public administration literature the reference is the citizens themselves. When the focus is the citizens themselves, empowerment is the government acknowledging that people should be involved in decisions regarding their own lives and acting accordingly. As Eran Vigoda (2002) notes, empowerment is seen when governments treat the citizens as customers or as partners, take their complaints seriously, encourage them to participate in open hearings, and, at more extreme levels, let them manage public institutions such as schools or housing estates. According to Eran Vigoda-Gadot and Shlomo Mizrahi (2008), this view has gained major attention in recent decades, and a rich discussion of public sector reforms has evolved, which has been heavily influenced by the new public management (NPM) doctrine that highlighted the role of citizens in governance and the meaning of customer orientation in public service.

Finally, the literature mentions empowerment in regard to subnational governments. When the focus of empowerment is on subnational government, the discussion is around "decentralization" or "devolution" of power, with the notion that local or regional decision makers know better what matches the needs and capabilities of their respective public. However, such empowerment is much debated as it produces little coherence in decision making and does not necessarily result in superior public policy. In addition, governments may empower social movements to carry out actions and legislation relevant to a specific social movement or to represent specific ideas in government deliberations. In recent years, the question of social movements and third-sector empowerment has intensified. Should government encourage more empowerment of those bodies and by so doing improve services to citizens? Or, could this be a sign of the hollow state that is practically incapable of meeting the expectations of the people with its own resources and bureaucratic agencies? This question remains moot.

Conclusion

Taking both the aggregative and integrative theories together, it is easier to understand the complex meaning of empowerment in public sector domains. Beyond the conventional meaning of engagement in organizational life, empowerment in the public sector has deeper and wider implications. It needs to refer to more stakeholders. It must consider the effect on larger sections of society, and it deals with forming and shaping the nature of our democratic values. Therefore, public sector empowerment has a widespread effect beyond the organizational arena. Its echo is heard far beyond the realms of public organization and affects the inner fabric of government and governance in modern societies.

Empowerment is thus a concept that is of much relevance in modern times to organizations, markets, governments, and people. Its generic managerial exploration diffuses into the public sector and receives additional interpretations that reflect the complexity of public organizations and the need to involve multiple stakeholders and subcontract with institutions on the federal, national, and local levels. Empowerment in the public sector means continuous interactions with the values of democracy and citizens' involvement and with the need to decentralize activities from state agencies to others. In the coming years, the public sector will undoubtedly make use of the ethos of empowerment to move forward with reforms and change and to improve services using self-mechanisms of public employees and citizens alike.

Dana R. Vashdi and Eran Vigoda-Gadot
University of Haifa
Haifa, Israel

See also Administration; Administration Theory

Further Readings

Bowen, D. E., & Lawler, III, E. E. (1995). Empowering service employees. *Sloan Management Review, 36*(4), 73–84.

Burns, D., Hambleton, R., & Hoggett, P. (1994). *The politics of decentralisation: Revitalizing local democracy*. Houndsmill, UK: Macmillan.
Thomas, K. W., & Velthouse, B. A. (1990). Cognitive elements of empowerment: An "interpretive" model of intrinsic task motivation. *Academy of Management Review, 15,* 666–681.
Vigoda, E. (2002). From responsiveness to collaboration: Governance, citizens, and the next generation of public administration. *Public Administration Review, 62*(5), 527–540.
Vigoda-Gadot, E., & Mizrahi, S. (2008). Public sector management and the democratic ethos: A 5-year study of key relationships in Israel. *Journal of Public Administration Research & Theory, 18,* 79–107.

English School in International Relations

See International Society

Environmental Issues

Environmental issues are concerned with human actions that affect the biosphere that humans and other species inhabit. Interest in the conservation of particular species, habitats, or landscapes, in particular localities or countries, dates back to the 19th century, but a concern with international environmental issues that transcend national boundaries may be dated back to the beginning of the 1970s. The European Conservation Year in 1970, the first Earth Day in 1970, and the United Nations (UN) Conference on the Human Environment held in Stockholm in 1972, which initiated the UN Environmental Programme, raised and focused the level of awareness and interest. A number of institutional developments also took place around this time. The Environmental Protection Agency in the United States was established in 1970, as was the Department of the Environment in the United Kingdom. In the European Community, the 1972 Paris summit of heads of government called on the Commission to draw up an environmental policy and set up a directorate responsible for environmental protection. The Community published its first Environmental Action Programme in 1973. New nongovernmental organizations with a broader agenda than older conservationist bodies and with a greater reliance on methods of direct action were founded around this time, the most important being Friends of the Earth and Greenpeace, both in 1971.

Why did this new interest in environmental issues emerge at this time? A more affluent and better educated society was becoming less interested in prosperity per se and more interested in a broader definition of the quality of life. Rachel Carson's best-selling book *Silent Spring* published in 1962, which dealt with the effects of pesticide use on the countryside, raised public awareness of the fact that modern technology had costs as well as benefits. The fragility, vulnerability, and beauty of the Earth were emphasized by the first pictures of the planet taken from deep space in 1968. This initial period of the environmental debate was also influenced by a concern about world population growth and its impact on scarce natural resources. Incidents such as the mercury poisoning at Minamata Bay in Japan in 1959 showed that pollution could have serious public health consequences, a lesson reinforced at the end of the 1970s by the Love Canal incident in the United States where leaking toxic waste affected the health of children, among others.

This early phase of the environmental debate was characterized by a more integrated ecological approach to the problems being encountered, in which the ecosphere was conceived as a whole, with changes in one part of the system seen as having effects elsewhere. This approach was exemplified by the 1972 Club of Rome *Limits to Growth* report. The analysis carried out for this report by researchers at the Massachusetts Institute of Technology (MIT) used computer modeling techniques to explore the relationships between a number of variables such as industrialization, resource depletion, pollution, food production, and population. Seen in retrospect, the results were far too pessimistic and highlight the dangers of projecting current trends into the future, but the report was highly influential at the time and encouraged consideration of the environment as an interacting system rather than just particular pollutants that had been the focus of earlier ad hoc policy interventions.

As policy measures developed, the debate became much more fragmented. It became focused on particular pieces of legislation concerned with various forms of pollution (air, water, and soil) and with habitats and biodiversity. One of the reasons for this fragmentation was that devising effective policy measures created a necessity for very specific policy expertise that was not readily transferable from one arena to another. An expert on river pollution might be of little help in tackling problems of marine pollution and certainly could not contribute very much to the debate on substratospheric air pollution. One risk associated with this more fragmented approach was an emphasis on legislative outputs rather than a focus on their likely impact on outcomes in terms of the state of the environment. The development of the climate change debate has given a central organizing narrative to environmental policy once again but perhaps at the expense of subordinating other aspects of environmental policy to this central goal. Thus, the protection of biodiversity may be viewed through the lens of the effects of climate change.

The Internationalization of Environmental Issues

Initial environmental policy responses were at a national or European Union (EU) level. It was evident that many environmental problems crossed national boundaries. Within Europe, examples included the pollution of major rivers that passed through a number of countries such as the Rhine or the effect of acid rain from British power stations on Norwegian forests. However, the problem of the thinning of the ozone layer could not be tackled at a regional level but required effective global collective action. By the 1980s, the threat from ozone-depleting chemicals was generally accepted. There was evidence of substantial thinning of the ozone layer in the polar regions. Associated risks included an increased incidence of skin cancer, crop failure, and a possible contribution to global warming. The 1985 Vienna Convention provided for controls on the production of ozone-depleting chemicals—chlorofluorcarbons (CFCs). The 1987 Montreal Protocol froze production of numerous substances believed to be responsible for ozone depletion, to be followed by a 50% reduction by the end of the century. The 1989 London Amendment agreed to phase out the most dangerous CFCs.

These agreements were an example of successful, effective global collective action to tackle an environmental problem, but there were a number of specific conditions that facilitated success. A single major cause of the problem was identified and accepted, which was not the case for climate change. This made it relatively easy to focus on what had to be done. Substitutes were available for the withdrawn products and were relatively inexpensive. Action did not require cuts in industrial production or economic growth.

New Concepts in the Debate

The debate on environmental issues was being driven increasingly by international reports and events, even if the required policies needed action at national and local levels, often supported by increasingly successful Green parties. A particularly influential report was produced by the UN World Commission on Environment and Development on *Our Common Future*, often referred to as the Brundtland Report after its chair. The report introduced the concept of "sustainable development," which focused on intergenerational justice. It was defined as meeting "the needs of the present without compromising the ability of future generations to meet their own needs." It became a key defining objective of many environmental policies.

Another concept that became increasingly influential was that of the "precautionary principle." This originally emerged in Germany as the *Vorsorgeprinzip* or principle of precaution. Although it initially had a number of definitions, it became increasingly influential in international discourse. The principle was first incorporated in an international treaty in the 1987 Montreal Protocol and was also reflected in the 1992 Rio Declaration on Environment and Development. At the center of the concept is the idea of taking anticipatory action in the absence of complete proof of harm when there is scientific uncertainty about causal links. The 1998 Wingspread Statement produced by a group of lawyers, scientists, and environmental activists summarizes the principle in this way: "When an activity raises threats to the environment or human health, precautionary measures should be taken even if some causes and effect relationships are not fully established scientifically." A communication from the European Commission on the precautionary principle in 2001 emphasized that any measures based on it should be proportional to the chosen level of protection, nondiscriminatory in their application, and consistent with similar measures already taken.

The assumption that there was a zero-sum trade-off between environmental protection and economic growth was challenged by the concept of "ecological modernization." Following the Brundtland Report, it was argued that a high level of environmental protection was essential to the success of a modern economy. Increasing energy efficiency, reducing pollution, and using waste products could boost the profits of firms and the success of the economy as a whole. The environmental technology industry could itself make an important contribution to economic success, particularly for the first companies to enter the market, as has happened in Germany with regard to solar technology. Politicians began to refer to the benefits of "green-collar" jobs. One consequence of this new school of thinking was less emphasis on "command and control" forms of environmental regulation and more emphasis on a range of "new environmental policy instruments," such as eco audits, voluntary agreements, and labeling.

Climate Change

Climate change refers to the warming of the planet as the result of human-made greenhouse gases, principally carbon dioxide, but also with significant impacts from gases such as methane, nitrous oxide, and halocarbons. Considerable uncertainty attaches to the consequences of these changes in the atmosphere, but they are likely to include a rise in sea levels, changes in weather patterns, and more extreme weather events. There is a global distributional issue attached to climate change as those most likely to lose out as a result are poor low-lying countries that have low emissions themselves and countries in sub-Saharan Africa that may experience an increased incidence of drought, affecting their agriculture.

Devising an effective set of policy solutions to the challenge of climate change is difficult. One is dealing not only with considerable uncertainty but with future uncertainty. The problem is caused by everyone and affects everyone on the planet, so there is an evident public issue that needs to be tackled, but there are also collective action and free-rider problems. The scale of international cooperation required is unprecedented. Much of business was initially opposed to any action being taken, trying to argue that climate change was not occurring, or if it was, it was not caused by human agency. However, being a "climate change denier" has become a less tenable and respectable position, while emissions trading represents a policy approach to the problem that business generally finds acceptable. The 2006 Stern Report highlighted the economic costs of inaction. The United States refused to sign the 1997 Kyoto protocol, but both candidates in the 2008 presidential election supported action on climate change. Emerging countries such as China and India objected to having to reduce their emissions when the problem had been caused by countries that developed before them. However, even if an effective international agreement can be reached, greenhouse gas emissions result from billions of independent decisions by households and firms. The policy instruments available to governments, such as information, incentives, and regulations, may prove to be inadequate.

Wyn Grant
University of Warwick
Coventry, United Kingdom

See also Environmental Policy; Green Parties; Postmaterialism; Social Movements

Further Readings

Helm, D. (Ed.). (2005). *Climate-change policy*. Oxford, UK: Oxford University Press.

McCormick, J. (2001). *Environmental policy in the European Union.* Basingstoke, UK: Palgrave Macmillan.

United Nations World Commission on Environment and Development. (1987). *Report of the World Commission on Environment and Development: Our common future.* Geneva: United Nations. Retrieved December 10, 2010, from http://www.un-documents.net/wced-ocf.htm

Vig, N. J., & Kraft, M. E. (Eds.). (2005). *Environmental policy: New directions for the 21st century.* Washington, DC: CQ Press.

Weale, A. (1992). *The new politics of pollution.* Manchester, UK: Manchester University Press.
Wingspread Conference. (1998, January 23–25). *Wingspread statement on the precautionary principle.* Drafted at the Wingspread Conference Center, Racine, WI. Retrieved November 1, 2010, from http://www.gdrc.org/u-gov/precaution-3.html

ENVIRONMENTAL POLICY

For as long as humans have existed on planet Earth, they have sought to alter it to suit their purposes. In its most basic sense, environmental policy seeks to govern the relationship between humans and their natural environment. A relatively new area of state activity, it has grown enormously, particularly since the 1960s. Indeed, its spectacular rate and extent of growth has been such that it is no longer a small and fairly discrete area of policy but one that increasingly intrudes into virtually all other policy areas. At its core is an identifiable "environmental state" comprising specific ministries, agencies, and organizations whose mandate is to secure environmental improvements. Underpinning environmental politics is the tense relationship between this state and its opposite numbers representing the social and economic realms. As the following sections show, this inherent tension can be better understood first through tracing the historical evolution of environmental policy, then outlining its distinctive characteristics, and finally by identifying its underlying rationales. While this overview would suggest that environmental policy development has resulted in a great deal of state building, the overall impacts on the state of the environment are not so clear.

The Main Phases of Environmental Policy Making

One way to understand environmental policy is to trace its historical evolution. Modern environmental thought dates back to the preindustrial period. Before the 1960s, environmental policy making was primarily geared toward protecting human health from pollution and establishing designated areas of green space for leisure activities. These ends were mainly achieved by limiting pollution from point sources such as factories and establishing protected areas and national parks.

In a second phase broadly encompassing the 1960s and 1970s, environmental policy really took off and environmental pressure groups boomed. It is often said that the first evocative images of the Earth transmitted from deep space in 1968 were what really catalyzed public concern for what became known as "the environment." The environmental state was born in this period as many jurisdictions created environmental ministries and agencies. However, these responses remained isolated, uncoordinated, and generally reactive. Policies adopted were mainly regulatory in nature, specifying process and emission standards; many were very poorly implemented.

In the 1980s and 1990s, environmental policy entered a third phase. The sector witnessed a huge expansion in the scale of the new environmental states, which sought to address problems in a more preventive manner at source. Old ways of thinking (such as "pollution control") duly gave way to new ones ("ecological modernization" and "sustainable development"). Rather than promoting end-of-pipe standards, these sought to embed environmental concerns within wider systems of human production and consumption. More specifically, they supplemented "command and control" regulations with new environmental policy instruments such as eco-taxes, voluntary agreements, and product-labeling devices. The new environmental states also began seeking innovative ways to cooperate on addressing cross-border pollution and resource overexploitation. A number of international environmental agreements mushroomed in this phase, and supranational bodies such as the European Union began to assume an ever-greater role in environmental problem solving, as shown by Andrew Jordan (2005).

Environmental policy in the industrialized world is currently in a fourth phase, roughly covering the period since 2000. This phase has witnessed another set of policy developments as the implications of environmental policy have ramified into ever more cognate policy domains. In this phase, it has become even more obvious that environmental problems arise not from single-point sources such as factories and power stations but from countless everyday social practices such as travelling, shopping, and even using the Internet. State action, by implication, is a necessary precondition for progress, but society too must play its part in nurturing and delivering environmental policy.

The issue that most clearly encapsulates these challenges is that of climate change—arguably *the* greatest environmental threat the world has faced (Andrew Jordan, Dave Huitema, Harro van Asselt, Tim Rayner, & Frans Berkhout, 2010). Crucially, it links into many other social and economic challenges such as poverty reduction, energy insecurity, and massive biodiversity loss and is stubbornly resistant to simple "techno-fixes." Climate change cannot, in short, be left entirely to environmental policymakers, but it demands long-term and socially transformative adjustments—a systematic "carbon revolution" no less—which mirrors the Industrial Revolution of the 19th century.

It is hugely important to be aware of the differences between these phases while at the same time remaining cognizant of the underlying continuities. The best way to appreciate the importance of both is to explore what distinguishes the environment from other policy areas.

What Is Distinctive About Environmental Policy?

Environmental quality exemplifies what economists term a *nonexcludable* or *"public" good*. In other words, the benefits of higher environmental quality often accrue to the wider public rather than to specific individuals. By contrast, the costs of protection (e.g., the fitting of abatement equipment) tend to be borne by specific groups (principally those that pollute or who are prevented from developing their land). Consequently, there is a basic asymmetry between those that pollute (in the initial phases noted above, generally a minority) and those (the wider public) suffering the consequences. Polluters have strong incentives to unite and fight to protect their "rights" to pollute and have therefore tended to enjoy advantages over those enduring its disbenefits. The latter are often too widely dispersed or individually suffer too little to mobilize into a coherent group.

In the third and fourth phases, this asymmetry has arguably become less stark as mass consumption has gradually eclipsed dirty "point sources" as the main driver of environmental change. Nonetheless, the state's role is still to mediate between those supporting and those opposing development. Because environmental damage has its origins in otherwise socially legitimate activities such as driving, using electricity, and farming, the state is routinely called on to limit any damage imposed by one section on another. Initially, this was achieved by regulating polluting activities. As regulating is a technically demanding process, the state found itself relying heavily on polluters and scientists for information and, eventually, political support. Environmentalists routinely claimed that this skewed regulation was in favor of producer interests. In Phases 3 and 4, the state has therefore come under pressure to explore alternative instruments to promote more sustainable forms of mass consumption (W. Neil Adger & Andrew Jordan, 2009).

Environmental policy tends to be heavily reliant on science, both to establish the extent of environmental problems and to determine how to resolve them. However, science is rarely capable of delivering clear-cut advice to policymakers in this regard. On the contrary, different interests, including governments, tend to choose and interpret scientific findings selectively in order to justify their preconceived views. These tensions are most acute when policymakers are required to go "beyond science" and address problems, such as climate change, whose causes and consequences are not precisely understood.

Environmental problems are highly expansive—they spill across political borders and create problems extending well into the future. The precise spatial and temporal distribution of costs and benefits produces alliances of self-interest, which then battle to determine the scope and content of policy. This is particularly visible in cross-national settings. International environmental policy burgeoned in the third and fourth phases noted above.

Balancing these various factors is the core policy challenge in the environmental policy sector. According to widely adopted interpretations of sustainable development (see above), policymakers should address the development needs of current generations of humans without compromising the ability of future generations to meet theirs. While there is general agreement that sustainable development is preferential to unsustainable development, moving from the latter to the former raises difficult questions centered on justice and morality, to say nothing of due democratic process.

Fifth, environmental policy is probably unique in the extent to which it cuts across traditional sectors of policy making. One great unmet challenge identified in the third phase has been how to integrate environmental concerns into and across all policy areas. The irony is that were this to happen, there would be no need for an identifiable "environmental" policy. For the time being, this seems an unlikely prospect: Although environmental arguments tend to find a more receptive audience in policy networks surrounding environment departments, they clash powerfully with producer-oriented networks found in the agriculture, transport, and energy sectors. Environmental policymakers are attracted to integration because it potentially helps tackle problems at source. As Andrew J. Jordan and Andrea Lenschow (2008) note, the danger, however, is that the sectors may only accept integration on their own terms or, worse still, take active steps to "reverse integrate" environmental policy with their preferred (sectoral) objectives. Integration in other words, is a two-way, not a one-way, street.

Sixth, environmental policy is probably the only policy domain that seeks to actively protect nonhuman entities. But to what extent should moral significance be ascribed to plants and animals, especially if it involves limiting human choice and raising the economic costs of development? This has opened up many ethically charged debates between those adopting broadly anthropocentric positions and those favoring more ecocentric ones. The former regards nature as primarily a source of resources and amenity; the latter holds that the nonhuman world should enjoy much greater (and in some cases equivalent) moral significance. Although there is plenty of evidence to suggest that people are more ecocentric than they were in the 1960s, many (if not all) contemporary environmental conflicts are still driven by the clash between these two worldviews.

What Drives Environmental Policy?

Ultimately, the main source of environmental policy making is environmentalism, a broad social movement that emerged in the late 1960s and loudly demanded that "quality of life" issues should receive more attention. In the past, environmental concern has tended to exhibit a cyclical pattern, with particularly pronounced peaks in support in the late 1960s and the late 1980s. Closer scrutiny reveals that these short-term "pulses" often coincided with periods of economic growth and social introspection. Conversely, concern has tended to tail off during economic recession. It is not normally as pronounced in poorer sections of Western society or in developing countries, although academic views differ on this matter. Looked at over longer time periods, an underlying effect is evident—each policy phase has provided the foundations for the next. During the third and fourth phases in particular, environmental policymakers sought new ways to institutionalize environmental concern and escape what Anthony Downs termed the episodic "issue attention cycle."

Environmental Policy: Retrospect and Prospect

Environmental policy addresses the complex and reciprocal links between the underlying structure of modern economies and societies and the quality of their natural environment. As such, it is probably unique in its breadth and its political complexity. Since the 1960s, a new "environmental state" has evolved to produce and implement it, paralleling the emergence of the welfare state in the 19th and 20th centuries. Both represent lagged and socially mediated responses to the negative effects of industrialization. And naturally, both are highly fertile grounds for political science and public policy research. But what has all this activity produced? While environmental policy has witnessed a great deal of state building, its overall impacts are still far from clear. For sure, the most damaging forms of air and water pollution have largely been contained in the most industrialized states. Yet efforts to address the creeping decline in ecosystem health reported by the Millennium Ecosystem Assessment (2005) as well as to stabilize Earth's climatic system have been continually overwhelmed by the ever-growing scale and speed of human activity. Consequently, environmental policy remains, at best, a work in progress.

Andrew Jordan and David Benson
University of East Anglia
Norwich, United Kingdom

See also Environmental Issues; Environmental Security Studies; Europeanization of Policy; Green Parties; Public Goods; Sustainable Development

Further Readings

Adger, W. N., & Jordan, A. J. (Eds.). (2009). *Governing sustainability*. Cambridge, UK: Cambridge University Press.

Dobson, A. (2007). *Green political thought* (4th ed.). London: Routledge.

Goudie, A. (2005). *The human impact on the natural environment* (6th ed.). Chichester, UK: Blackwell.

Jordan, A. J. (Ed.). (2005). *Environmental policy in the European Union* (2nd ed.). London: Earthscan.

Jordan, A. J., Huitema, D., van Asselt, H., Rayner, T., & Berkhout, F. (Eds.). (2010). *Climate change policy in the European Union*. Cambridge, UK: Cambridge University Press.

Jordan, A. J., & Lenschow, A. (Eds.). (2008). *Innovation in environmental policy?* Cheltenham, UK: Edward Elgar.

Millennium Ecosystem Assessment. (2005). *Ecosystems and human wellbeing*. Washington, DC: Island Press.

Environmental Security Studies

Environmental security studies analyze both the confluence of foreign policy and environmental security, particularly in the aftermath of the Cold War, and the linkages between security cooperation and global environmental risks. They result from an opening in the fields of foreign policy, defense studies, and security cooperation, wherein researchers have tended to move away from a narrowly defined military and strategic understanding of threat, vulnerability, risk, and national interest. In a post–Cold War era characterized by growing transnational flows, emerging environmental movements, and a global political process of ecological awareness building, the changes produced in the regional and global environment have become a critical area on the security agenda. However, the term *environmental security* has acquired multiple meanings and has gone through contentious academic and policy debates about how environment and security can be connected. Neither scientific programs nor policy agendas have so far been able to produce a commonly agreed definition of environmental security, which also shows the academic relevance and the policy-building interest of continuous reflection in this research field. However, the same goes for the concept of security itself; as Edward Page (2000) affirms, environmental security appears to be a clear example of a concept that is prone to endless, irresolvable dispute regarding its meaning and policy application. This entry discusses how concerns with environmental security have evolved since the Cold War era and then examines its place in foreign policy today.

Origin and Evolution of Concerns With Environmental Security

In the 1960s, the first efforts were directed toward articulating issues about the environment and security. They were generally concerned with the impact produced by humans on the security of nature, animals, and plants and also with the links between development and quality of life. These first interpretations of environmental security have shed light on the misuse of pesticides and fertilizers and the threats to the environment associated with nuclear energy. As in the cases of Rachel Carson's book *Silent Spring* (1962), Aldous Huxley's article "The Politics of Ecology: The Question of Survival" (1963), and Kenneth E. Boulding's notion of spaceship Earth (1966), these perspectives have emphasized the need to secure the integrity of ecosystems in order to ensure mankind's survival. Mercury poisoning associated with the Chisso-Minamata disease (1959) and the Torrey Canyon shipwreck off the western coast of Cornwall in England (1967), among other environmental disasters, have played a key role in this context.

In the 1970s, other readings analyzing the environmental effects of war and refugee movements developed, and research also tackled issues related to the link between the environment and development. This was the case with the "Only One Earth" report written by Barbara Ward and René Dubos, Edward Goldsmith's "Blueprint for Survival," and the "Limits to Growth Meadows" report, all of them published in 1972. They all expressed an overreaction to some of the supposed dangers of environmental degradation to human security. The United Nations (UN) Conference on Human Environment (Stockholm, 1972), despite the fact that it ended with a largely dead minimalist consensus among states on 26 principles and 109 nonbinding recommendations, has paved the way for future environmental cooperation, led to the establishment of global and regional environmental monitoring networks, and allowed the creation of the UN Environmental Program in Nairobi. In its final declaration, there is not a single reference to environmental security, although its Principle 21 states that "States have . . . the responsibility to ensure that activities within their jurisdiction or control do not cause damage to the environment of other states or of areas beyond the limits of national jurisdiction." Herein resides the idea of environmental threats to the national interest of a state, which should eventually produce a change in the conception of foreign and defense politics.

Since the 1992 UN Conference on Environment and Development held in Rio de Janeiro, most of the academic debate has been conceptual, leading to the creation of several schools of thought. Some researchers have analyzed the significance of environmental stress as a contributing factor to conflict in many developing countries, generally in an attempt to relate environmental scarcity (environmental degradation or the depletion of renewable resources) to violent conflict. Research has also developed on environmental change produced by human activities and conflicts and shifted to examining how environmental stress may contribute to conflict in combination with other relevant factors. Since 1989, for instance, Thomas Homer-Dixon and his team have produced a significant body of research on the connections between large-scale human-induced environmental stress and mass violence in developing countries. This team of researchers has analyzed the links between water scarcity and conflict, rapid urbanization and urban violence, and environmentally induced migration and ethnic violence, among other issues, having conducted case studies of environmental stress and violence in Rwanda, South Africa, Bangladesh, India, Pakistan, China, Nicaragua, Mexico, and so on.

More recently, research has highlighted the importance of disputes arising from access to natural resources such as water, oil, or minerals for strategic purposes. Some researchers call them *new wars*. It is clear that for some, what needs to be secured remains predominantly the survival of the state, and in this sense environmental insecurity can be read as synonymous with environmental threats to the nation-state (e.g., military operations may have an impact on the environment or the environment may be manipulated or contaminated as a target during wars). For other researchers, environmental security has a much broader scope, including the security of individuals and communities and the security of nature itself. Some more critical researchers associate environmental security with issues of inequality in North–South relations, social class, and revenue distribution. To them, the central research questions are "Whose security is at stake?" and "Who defines the need to protect what and for whom?"

These theoretical debates also show clearly that the relationship between the environment and security is historically and politically complex and controversial and that many factors play a role in it. The cause-and-effect relationships between tensions and vulnerabilities are multidimensional, and the links between the various political and economic components may be direct or indirect. These vibrant discussions also reflect different concepts of nature and the very distinct social representations of what should get counted as part of the environment. One of the dangers of some approaches is that they often dichotomize human beings and nature. On the one hand, deep ecology may often see nature as a subject in itself—a subject that requires protection and for which technical norms should be set up. On the other hand, anthropocentric visions tend to promote human interests over those of the environment. If one understands the notion of environment as including humans, then the way one defines environmental problems alters, and one may come to reformulate the concept of environmental security in terms of broader human security. What constitutes the environment is, in this respect, a key conceptual matter for our understanding of environmental security.

Environmental Security and Foreign Policy

Foreign policy has traditionally referred to a series of actions followed by one nation to deal with another nation, another region, or an international issue in a multilateral setting. In this connection, a state's foreign policy would be based on values (democracy, promotion of peace and human rights, rule of law, etc.) and national interest and would usually be aimed at preserving the state's economic and political goals abroad as well as its position in the world. Foreign policy analysis has classically oriented its lenses to the actions of national governments, principally those related to diplomatic or military issues, thus reflecting an overly simplified view of a world justified in the name of a well-established sense of national interest and threats to national security. Today, the media, the corporate sector, lobbyists, think tanks, environmental and human rights nongovernmental organizations, ethnic groups, and social movements, among other actors, play an increasingly important role in trying to influence foreign policy in many countries. Furthermore, with the end of the Cold War, the in/out divide has lost its absolute relevance, and there has been a keener interest in domestic policy issues, such as energy policies, ecological and health issues, the use of chemicals in agriculture, and the impact of genetically modified organisms in the environment. These developments have deeply changed the analytic field of foreign policy, its agendas, decision-making processes, and the definition of the social practices of its stakeholders. Taking into account environmental security is but another intellectual and political challenge put forward to both the epistemic community and the world of practitioners of foreign policy.

In fact, Braden Allenby (2000) suggests that environmental security may be an important evolution in the definition of what constitutes the national interest of states in international policy systems. Because the previous apparent stability generated by the ideological confrontation between capitalism and communism has broken down, global environmental changes and the need to promote environmental security worldwide potentially affect both foreign policy and security agendas more easily. Environmental issues can no longer be thought of as merely subsidiary components of industrial, social, political, and economic systems. There is a movement of environmental problems from an "overhead" to a "strategic" position on the global agenda. This movement from "periphery" to the "center" occurs at many different scales within firms, industrial sectors, society itself, and governments. They may differ in detail and timetable but have many fundamental similarities in each case and across regions—both North–South and East–West.

This happens also because there is an important constellation of environmental issues (energy, access to natural resources, climate change and sea-level increases, biodiversity-related wealth, dangerous waste management, etc.) that support the fabric of a country's security and its territorial integrity. They may also threaten and be of most concern for foreign policy analysts and decision makers. They challenge two comfortable classical assumptions that have been at the basis of traditional views of foreign policy: (1) that the nation-state is relatively absolute in the definition of its sovereign security interests and that (2) only interstate relations (particularly within Cold War conflicts between capitalism and communism) define global geopolitics. From a state-centric perspective, security is national; thus, environmental threats to national security should be dealt with by national authorities.

The integration of the notion of environmental security (and insecurity) into the field of foreign policy makes these assumptions, at least in their absolute form, much less valid.

It is in this connection that the Worldwatch Institute had already demanded, in 1977, a broadening of the sense of national security. Lester R. Brown, in "Redefining National Security," had reaffirmed that threats to countries and societies should stem less from relations with nations and more from relations with nature. Jessica Mathews makes the same argument in her 1989 article on "Redefining Security." Even within the more traditional realist scholarship, there is an emphasis on this link; for example, George Kennan claimed (in 1985) that the degradation of the environment is one of two threats facing mankind, the other being military and terrorist attacks. Extreme versions of realism, such as the geopolitical theories of major-general Karl Haushofer, look at the security implications of strategic raw materials. Both German and Japanese expansion in the 1930s would partly have been a search for raw materials. Some researchers see President George Bush's intervention in Iraq in 2003 as an attempt to secure the oil resources of the Middle East.

Indeed, the foreign policy of the United States of America has begun to evolve in response to these recent environmental security challenges, in part by recognizing the need to manage a new set of issues, broadening the generally accepted and cozy concept of security as part of a larger national security mission. In a speech at Stanford University in 1996, the then U.S. Secretary of State Warren Christopher had explicitly recognized the need to include additional dimensions in American foreign policy, noting that "our Administration has recognized from the beginning that our ability to advance our global interests is inextricably linked to how we manage the Earth's natural resources," and the great relevance to "contend with the vast new danger posed to our national interests by damage to the environment and resulting global and regional instability." Irrespective of the rhetorical dimension of such a statement, the U.S. secretary of state then concluded on a strong note: "That is why we are determined to put environmental issues where they belong: in the mainstream of American foreign policy." In the case of Japan, in his 1980 "Report on Comprehensive National Security," the then prime minister of Japan asserted that security would encompass economic vulnerabilities, natural disasters, environmental degradation, and ecological imbalances.

These developments confirm what Thomas Homer-Dixon and Jay Blitt, in their 1998 book *Ecoviolence*, had pointed to as their key findings in the analysis of linkages between environmental scarcity and foreign policy:

1. Scarcities of renewable resources produce civil violence, instability, and conflict.
2. The degradation and depletion of renewable resources causes environmental scarcity.
3. Powerful groups capture valuable resources, while marginal groups migrate to ecologically sensitive areas.
4. Environmental scarcity constrains economic development and produces migration.
5. Existing distinctions between social groups may be sharpened by environmental scarcity.
6. Environmental scarcity weakens governmental institutions and states.
7. Environmental scarcity can also cause ethnic conflicts, insurgencies, and coups d'état.
8. The international community can be indirectly affected by these conflicts generated by environmental scarcity.

These discoveries support the assumptions of political ecology researchers who argue that environmental degradation often leads to economic underdevelopment, which can in turn diminish political stability in countries and societies in the South.

Global Environmental Risks and Security Cooperation

There is no doubt that today's human activities constitute a significant and frequently destructive feature of Earth's ecosystem. Interpretations may vary, but UN sources suggest that around 40% of the entire natural photosynthetic product of the biosphere is appropriated by human beings. It is clear that the world faces a global environmental crisis and, inseparable from this, a coincidental crisis of growing social inequality and poverty. Unprecedented environmental and social changes pose huge challenges to the interstate system, one key question being if and how a global set of local sovereign states would be capable of saving the biosphere from environmental degradation and pollution. Many signs indicate a need for society's cross-sectoral attention to the environment as an underlying security issue. This sentiment had echoed through the literature on environmental and ecological problems and also through the opening words of the Brundtland Report in 1987: "The Earth is one but the world is not. We all depend on our biosphere for sustaining our lives. Yet each community, each country, strives for survival and prosperity with little regard for its impact on others" (World Commission on Environment and Development, 1987).

Although security tends to be defined only from a military perspective and is thus not open to environmental concerns, since the end of the 1980s some debate has been devoted to global environmental risks and to the threats that they represent to collective security. Al Gore, in his 1992 book, *Earth in the Balance*, had pointed to global risks such as the production and trade of highly toxic chemicals, the loss of biodiversity, ozone depletion, climate change, marine degradation, and desertification and deforestation. The environment, peace, and security chapter of *Our Common Future* was not subject to much negotiation at the 1992 United Nations Conference on Environment and Development in Rio de Janeiro.

Developing countries did not endorse a global dialogue on environmental issues within the context of conflict and security, reacting negatively to any multilateral treatment of environmental security proposals within UN forums. Coalitions of developing nations, mainly the Group of 77, perceived the security frame as a Pandora's box that, once opened, could formally menace their claims of sovereign control over their resources. Moreover, leading developing nations (Brazil, China, India, and Malaysia, among others) tend to consider that predominantly Northern geopolitical concerns about the impacts of environmental conflicts on national sovereignty and the destruction of the global commons associated with comprehensive security have largely dominated environmental security negotiations and policy making to date. Their criticisms focus on the failure of this security approach to address the causes of inequality and injustice that underpin poverty and environmental degradation at local and global levels. Representatives from developing nations support that global change has more impact on those least able to protect themselves from its effects, such as the poor and coastal communities at risk from severe flooding and climate change spillovers.

Therefore, the world of interstate relations has not welcomed the ideal of environmental responsibility as another normative principle in international relations, arguing that it would justify breaches in the state's sovereignty and legitimize foreign interventions by stronger countries. That is why some authors such as Norman Myers tend to defend the individual level of analysis; that is, security should apply mostly at the level of the individual citizen as a basis of political stability. In his view, security should not be a function of stable macrosocial processes (such as sociopolitical, economic, or even military systems); rather, it should be seen as a goal to be achieved at the microlevel of human existence. Another type of definition is provided by ecological security approaches, which tend to consider that environmental degradation, by undermining mankind's means of subsistence, threatens the security of ecosystems. Ecological security takes into account the need to provide for the physical circumstances and conditions of the needs of a community without diminishing natural stocks. Other approaches—for instance, the one supported by the United Nations Development Program—are inclined to privilege human security as an overarching concept. As a consequence, environmental security and human rights would be inextricably linked. In this case, *security* should become an all-encompassing term relating to the social, economic, political, and ecological wellbeing of individual human beings.

Irrespective of the level of analysis, all these perspectives agree that environmental threats pose long-term dangers to security and that the causes of environmental insecurity are related to population growth, degradation, overuse and abuse of resources, depletion of finite global resources, transborder effects of environmental degradation, or the energetic impact of unsustainable consumption patterns. Apart from the intrinsic shortcomings associated with the interstate system to deal with global environmental security issues, another major obstacle that may stand in the way of effective global governance mechanisms and structures relates to the extraordinary power of transnational corporations. Of the 100 largest economies in the world, 51 are corporations; and the top 500 corporations now control almost two thirds of world trade.

The literature on environmental security is very much shaped by the assumption that the behavior of states is determined by the structure of power relations in international relations. Particularly in the 1990s, this literature has also been under the influence of a credo that no regulatory mechanisms should guide or control corporate behavior. In this sense, security would be restricted to the need to protect a population and a territory against foreign threats while at the same time defending the interests of the state. Questions related to scope and content would be withdrawn from the policy debate, but some scholars insist on retaining them: What is the nature of the societies, organizations, and institutions whose security matters from the point of view of security discourse? What environmental phenomena put them at risk and render them insecure? Who is threatened by environmental insecurity?

According to Steven Ney (1999), it is necessary to recall that evaluating different models, or conceptual maps, of environmental security is not a straightforward task. Using Graham Allison's terminology, models can be considered *perceptual lenses* that allow us to make sense of the world: A theory is a way of seeing as well as a way of not seeing. As such, models are tools with which policy actors may construct arguments in the course of policy negotiation and deliberation. Models can have a pragmatic impact on policy making, and by extension they may be designed solely to affect the policy process. They may perform as well as inform. That is why it does not make sense to assess models in isolation of their argumentative contexts and concrete policy settings in which they are to be employed. As constructivist and critical researchers argue, the often implicit normative assumptions that underpin the different approaches to environmental security should also be analyzed. For instance, one of the explanatory powers of many environmental security approaches is based on the assumption that increased population growth inevitably leads to environmental degradation. By the same token, however, this argument does not take into consideration how individual and social consumption patterns are structured and how some of them may have a stronger impact on the environment. It is true that the various theoretical approaches to environmental security contribute to promoting global environmental issues into policy agendas worldwide, but the question that environmental security studies should also pose relates to how these issues are being brought into the agendas.

Carlos R. S. Milani
Universidade Federal da Bahia
Salvador, Bahia, Brazil

See also Environmental Issues; Environmental Policy; National Interest; Security and Defense Policy; Security Cooperation

Further Readings

Allenby, B. R. (2000). Environmental security: Concept and implementation. *International Political Science Review, 21*(1), 5–21.

Barnett, J. (2001). *The meaning of environmental security: Ecological politics and policy in the new security era*. London, New York: Zed Books.

Brown, L. R. (1977, October). *Redefining national security* (Worldwatch Paper No. 14). Washington, DC: Worldwatch.

Dabelko, G. D. (2008). An uncommon peace: Environment, development, and the global security agenda. *Environment, 50*(3), 32–45.

Eddy, E. (2004). Environmental security: Securing what for whom? *Social Alternatives, 3*(4), 23–28.

Foster, G. D. (2001). Environmental security: The search for strategic legitimacy. *Armed Forces & Society, 27*(3), 373–395.

Homer-Dixon, T. F. (1999). *Environment, scarcity, and violence*. Princeton, NJ: Princeton University Press.

Homer-Dixon, T. F., & Blitt, J. (Eds.). (1998). *Ecoviolence: Links among environment, population, and security*. Lanham, MD: Rowman & Littlefield.

Le-Prestre, P. G. (1993). La sécurité environnementale: Nouvel objectif de la politique de défense américaine? [Environmental security: New objective of the politics of American defense?]. *Relations internationales et stratégiques, 11*, 35–45.

Mathews, J. (1989, Spring). Redefining security. *Foreign Affairs, 68*(2), 162–177.

Myers, N. (1993). *Ultimate security: The environmental basis of political stability*. New York: W. W. Norton.

Ney, S. (1999). Environmental security: A critical overview. *Innovation* (Interdisciplinary Centre for Comparative Research in the Social Sciences), *12*(1), 7–30.

Page, E. (2000). Theorizing the link between environmental change and security. *RECIEL, 9*(1), 33–43.

Smouts, M.-C. (2001). Risque planétaire et sécurité environnementale [Planetary risk and environmental security]. *Esprit, 274*, 133–141.

United Nations Environmental Programme. (1972). *Report of the United Nations Conference on the Human Environment*. Retrieved November 25, 2010, from http://www.unep.org/Documents.Multilingual/Default.asp?documentid=97

World Commission on Environment and Development. (1987). *Our common future* (The Bruntland Report). Oxford, UK: Oxford University Press.

Epistemic Communities

Epistemic communities are networks of knowledge-based communities with an authoritative claim to policy-relevant knowledge within their domain of expertise. Their members share knowledge about the causation of social or physical phenomena in an area for which they have a reputation for competence and a common set of normative beliefs about what actions will benefit human welfare in such a domain. In particular, they are a group of professionals, often from a number of different disciplines, who share the following set of characteristics:

1. *Shared principled beliefs:* Such beliefs provide a value-based rationale for social action of the members of the community.
2. *Shared causal beliefs or professional judgment:* Such beliefs provide analytic reasons and explanations of behavior, offering causal explanations for the multiple linkages between possible policy actions and desired outcomes.
3. *Common notions of validity—intersubjective, internally defined criteria for validating knowledge:* These allow community members to confidently differentiate between warranted and unwarranted claims about states of the world and policies to change those states.
4. *A common policy enterprise:* This entails tackling a set of practices associated with a central set of problems, presumably out of a conviction that human welfare will be enhanced as a consequence.

Epistemic communities is a concept invoked by constructivist scholars of international relations to focus analytic attention on the process by which states formulate interests and reconcile differences of interest. Epistemic communities are a principal channel by which consensual knowledge about causal understandings is applied to international policy coordination and by which states may learn through processes of international cooperation.

John Ruggie introduced the term *episteme*, borrowed in turn from Michel Foucault, to describe the overarching perspective through which political relationships are visualized and understood during historical eras. The 1992 winter issue of *International Organization* on "Knowledge, Power, and International Policy Coordination" focused on the actors responsible for articulating and aggregating knowledge-based understanding in areas of security, environment, and international political economy (IPE). It also developed the now standard four-element definition of an epistemic community provided above.

By analyzing epistemic communities, constructivist scholars gained leverage on understanding the processes of social construction and collective learning. If contemporary international relations is characterized as a setting of complexity and uncertainty, as argued by constructivist scholars—particularly under contemporary circumstances of complex interdependence, increasing globalization, and the emergence of new technical issues on the international agenda with which traditional decision makers are habitually unfamiliar—then state interests are often unknown or incompletely specified. International relations then becomes a matter of applying embedded and institutionalized beliefs about the nature of problems and the appropriate means of collective response rather than the process of resolving rationally formulated state preferences as argued by many current theorists of international relations. Changes in information processing are likely to follow well-publicized shocks or crises. Only at such times are decision makers likely to recognize major anomalies and pursue new policy patterns. During subsequent, less revolutionary periods, these new doctrines or orthodoxies assume the status of taken-for-granted assumptions or dogma, which persist until called into question again by external stimuli. Because of the disjointed equilibrium nature of policy change, an evolutionary focus on institutional learning and path dependence may provide an appropriate model by which to understand the international recognition and response to global change. Such a research program may provide a better understanding of factors that influence the introduction of new policy frames, collective understandings, or doctrines as well as illuminate mechanisms of lock-in and identify those factors that may influence the degree of irreversibility of national and collective actions.

Epistemic communities are one of the principal actors responsible for aggregating and articulating knowledge in terms of state interests for decision makers and disseminating those beliefs internationally. In a broader political context, epistemic communities provide one of the major channels by which overarching regime principles, norms, and rules are articulated for the international community and disseminated internationally. While epistemic communities are the principal agents responsible for articulating such principles, norms, and rules, the extent to which they become more deeply diffused and embedded internationally has to do with the political influence of epistemic community members: their ability to persuade others, their ability to consolidate bureaucratic influence in important institutional venues, and their ability to retain influence over time. State interests and decisions to deploy state power are thus identified subject to consensual knowledge.

Research on Epistemic Communities

The epistemic community literature is now nearly 20 years old. The epistemic communities concept was initially favorably received in international relations because it provided a means for focusing on the ideational component of politics and also allowed for agency in theorizing about governance and policy making. Studies of the European Union (EU) in particular have analyzed the role of various epistemic communities in shaping EU directives as well as in building a broader sense of European identity. Since the initial response to epistemic communities, analysts have developed a variety of clarifications, refined hypotheses, carried out further empirical work, and empirically confirmed the broad predicted social patterns associated with epistemic communities.

Initial critiques of epistemic communities called for a clearer theory of the state as well as a clearer metric by which epistemic communities could be recognized and consensus within an epistemic community could be measured.

Critical voices from the literature were sounded on social studies of knowledge regarding the degree of political autonomy enjoyed by epistemic communities and science-based arguments in general for public policy, questioning the political consequences of such ideas and also the potential implicit political bias in the research programs pursued by epistemic community members. They feared that by depoliticizing expertise and the consequences of expert-based advice, epistemic communities may become antidemocratic and antiparticipatory. However, such points were explicitly acknowledged in the initial formulation of the epistemic community research program—namely, the political nature of all policy debate. The epistemic community argument was that, normatively, epistemic communities ultimately provided more impartial advice than other modes of policy advice and, analytically, their study offered a clear causal pathway by which ideas came to inform political practices that was superior to conventional approaches to the study of politics, which are unable to provide credible explanations for how ideas influenced politics or the conditions under which ideas were likely to be influential.

Epistemic community members can be identified through interviews, from secondary literature (especially by journalists), and by cross-checking membership lists for international negotiations and consultations over time. Recurrent names are eligible candidates. Refining an identification of the consensual knowledge shared by the epistemic community members can be done by reading their publications (especially scientific ones with equations that force precision) and open-ended interviews. The extent to which they feel that they are members of a common community can be ascertained through interviews and through snowball techniques, in which prospective members identify others with whom they believe they share beliefs. The combination of self-identified traits and externally confirmed traits (such as the application of consensus approaches to truth) define the epistemic community. The causal mechanisms by which they exercise influence—learning—can effectively be pursued by process tracing, focused comparative case studies, counterfactuals, and alternative hypothesis testing.

Stronger hypotheses about the mechanisms, effects, and variation of epistemic communities' influence were also developed. Epistemic communities are likely to be found in substantive issues where methods from scientific disciplines have been applied to policy-oriented work and in countries with well-established institutional capacities for administration, and science and technology. Governments with such capacities are likely to see the need for the technical skills that epistemic community members command, and such professionals may be attracted to governmental service only when they believe that their policy enterprise can be advanced. Crises or widely publicized shocks are probably necessary precipitants of environmental regime creation, but crises alone are insufficient to explain how or which collective responses to a perceived joint problem are likely to develop.

Epistemic communities then help identify cause-and-effect relationships, elucidate interlinkages between problems, help define the consulting state's or organization's interests, and help formulate policy. Their aggregate effect depends on the extent to which their ideas become embedded in influential multilateral institutions more generally: powerful countries and international institutions and, possibly, nongovernmental organizations (NGOs), which will then deploy their own influence to disseminate the shared ideas enunciated within the epistemic community. Overall, learning will occur in the policy system as new policy-relevant knowledge is identified and applied to a common problem.

The initial reception of their ideas is likely to occur in rich democratic societies, where scientific capacity and the free flow of information within society are high. However, other countries may also be receptive to epistemic communities' ideas but not as quickly.

Treaties developed with a strong contribution from epistemic communities are likely to enter into force more rapidly than those developed through other political mechanisms, as the social influence of epistemic community members will accelerate ratification in domestic legislative venues.

Multiple mechanisms of diffusion are possible through various international channels at different political scales. For instance, in stratospheric ozone protection, the very regulatory scheme arose from atmospheric chemists, but the new technologies to satisfy those regulations emerged through the interplay of multinational corporations (MNCs, e.g., DuPont, ATOCHEM, and ICI), international institutions (the Montreal Ozone Protocol's technical advisory panels and the Montreal Ozone Fund), and NGOs (most notably, Greenpeace). The appropriate factors by which international institutions may induce other actors to accept the epistemic ideas will vary because of the political and economic conditions in the target country; for instance, advanced industrialized societies are more responsive to public education campaigns, whereas poor developing countries respond to material incentives from capacity building. The key analytic point, though, is that without the involvement of epistemic communities to set the agenda and impart new ideas, the political dynamics of collective action are likely to be far more conventional and lack the potential for reflective learning by policymakers.

Subsequent empirical work has found that epistemic community members have played a strong role in negotiating international environmental agreements. Members of the epistemic community that have dominated technical discussions in environmental regimes have subscribed to holistic ecological beliefs about the need for policy coordination subject to ecosystemic laws. Thus, they promote international environmental regimes that are grounded on policies that offer coherent plans for the management of entire ecosystems and are sensitive to interactions among environmental media (such as air and water), sources of pollution, and contending uses of the common property resource, rather than being limited to more traditional policies for managing discrete activities or physical resources spaces within fairly short-term time horizons. Epistemically informed collective action has a distinctively comprehensive form and is politically resilient, yielding patterns that are not associated with political dynamics involving any other sets of political actors. Moreover, epistemically informed treaties are more likely to be effective because the regulatory standards fit the behavior of the socio-ecosystem being managed.

Further studies of international institutions refined the characteristics of formal international organizations that are likely to absorb the lessons of epistemic communities and disseminate them. Design principles for international scientific panels were also identified.

The hypotheses about the distinctive patterns of epistemic collective action were confirmed by numerous focused comparative case studies of a wide array of multilateral environmental regimes. In IPE (international political economy), similar dynamics are associated with the popularization and dissemination of major economic doctrines, including the Washington Consensus, Keynesianism and structuralism, and development economics.

The ideational focus was absorbed into the broader constructivist research program developed in international relations and comparative politics, which looked at the role of beliefs and ideas in shaping state interests and practices, with epistemic communities serving as one of the mechanisms by which new ideas are developed and circulated. Analyzing epistemic communities continues to provide a key group of actors associated with a distinctive set of processes by which causal understandings shape actor interests and regime dynamic amid constructivists' broader focus on other ideational forces, including norms, linguistic usage, and the like.

More recent publications in the epistemic communities research program have been looking more carefully at the context in which epistemic communities arise and operate. Current works try to clarify the institutional factors that shape or amplify ideational consensus and dissemination and that influence the creation of relatively impartial and usable scientific advice, and they try to separate the causal mechanisms of social learning from other causal mechanisms that drive collective action. For instance, how can one explain which epistemic community will prevail and when? One could ask why environmental scientists have tended to prevail over economists in all environmental issues other than climate change. Additional research frontiers include the mechanisms of lock-in, the domestic dynamics that influence the authority, the identity and influence of epistemic communities, and the mechanisms by which decision makers learn more about the nature of the international political system as well as the management of isolated issues.

Peter M. Haas
University of Massachusetts Amherst
Amherst, Massachusetts, United States

See also Elitism; Governance; Intellectuals; Power

Further Readings

Adler, E. (2005). *Communitarian international relations.* London: Routledge.
Dunlop, C. (2000). Epistemic communities: A reply to toke. *Politics, 20*(3), 137–144.
Haas, P. M. (Ed.). (1997). *Knowledge, power and international policy coordination.* Columbia: University of South Carolina Press. (Original work published 1992)
Haas, P. M. (2001). Policy knowledge: Epistemic communities. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 11578–11586). Oxford, UK: Elsevier.
Haas, P. M. (2004). When does power listen to truth? *Journal of European Public Policy, 11*(4), 569–592.
Haas, P. M., & Haas, E. B. (1995). Learning to learn. *Global Governance, 1*(3), 255–285.
Haas, P. M., & Haas, E. B. (2002). Pragmatic constructivism and the study of international relations. *Millennium, 31*(3), 573–601.
Maasen, S., & Weingart, P. (Eds.). (2005). *What's new in scientific advice to politics?* Dordrecht, Netherlands: Springer.

Epistemological and Methodological Foundations

Understood literally, *method-ology* means the study and/or use of methods. This understanding implies the existence of more than one method and that the utility of a particular method can vary across contexts. For many practitioners in political science, the method of choice is determined by the number of observations: Students of relatively abundant political phenomena rely on statistical techniques, students of unique political phenomena use case study techniques, and those in the middle sample choose from a wide array of qualitative or small-*N* comparative approaches. This methodological rule of thumb is derived from an implicit hierarchy of methods, which is the product of a particular, and very influential, research tradition.

This entry examines the ontological and epistemological foundations of this research tradition to explain the reasoning behind its implicit hierarchy of methods and to show some of its inherent limitations. There are two objectives for this. The first objective is to broaden the reader's methodological horizons by showing how one's choice of methods should reflect the nature of the subject one studies as well as the way in which one can acquire knowledge about that subject. By examining these methodological foundations, political scientists can become more aware of both the potential and the shortcomings of the methods they use. Just as important, methodological reflection can encourage political scientists to use a broader range of methods. In recognizing how the usefulness of any particular method will depend on the nature of the subject under study as well as on the mode of explanation that the analyst deems most appropriate, political scientists may find it necessary to adapt their methods, or adopt new ones, as the nature of the political world (and political scientists' interest in it) evolves.

To stimulate this sort of reflection, this entry focuses on the philosophical underpinnings of the most common approaches. For empirically minded political scientists, these philosophical concerns are twofold: (1) they need to be aware of the nature of the things they study, and (2) they need to concern themselves with the various means by which they can secure reliable knowledge about those things. These two concerns provide the basic framework for this entry. The first section examines the way in which mainstream approaches depend on a particular understanding of the political world to produce the theories and generalizations that the political science profession admires. The most common methods are designed to exploit specific ontological contexts and epistemological standards; as a result, they do not always work as designed when transplanted to new and different contexts. The second section introduces an alternative ontological approach and its relationship to the same, commonplace methods. In short, the relationship among ontology, epistemology, and the methods is discussed in the first two sections that follow. The third section examines some of the most common difficulties facing researchers who are trying to secure reliable knowledge about that political world. In their attempt to secure and understand knowledge about the world, they inevitably impose their own ways of ordering it. Recognizing and correcting these inherent biases is the subject of the third section, which harks back to Francis Bacon for the task.

The Political World

Political science—as an academic discipline—has grown out of several fields, which vary significantly from one national tradition to the next. In some contexts, the weight of history has been strongest; in others, politics grew from the study of law; in still others, economics or philosophy were the host communities. As a result, modern political science can draw from a very deep methodological well, if only one is willing to tap into it. Throughout most of the history of European political thought, one can find a remarkable diversity and openness about how to address political affairs. The founding thinkers of the political science discipline—Plato and his student Aristotle—could not agree about the nature of the world they studied, the utility of different sources of knowledge about that world, or on any particular method for studying that world. After all, Plato was not especially interested in experience or empirical detail; he emphasized the importance of reason in discovering truths about an ideal world. Aristotle, by contrast, favored a more down-to-earth approach, where observation of events in the world and careful comparisons could show us true from false.

In these foundational thinkers, one can find an open discussion about the importance of linking the approach of study to the world of study. In the centuries since, there has been a broad consensus that the nature of the social and political world is relatively similar to that of the natural world. Indeed, this consensus is one reason why many methods developed for studying the natural world have successfully migrated to the social sciences, and it may be an important reason why political scientists spend so little time discussing ontological issues. As will be seen in the next section, this is not the only ontological perspective embraced by political scientists—but it remains the most dominant. Most political scientists view the political world in a way that is strongly influenced by a long and illustrious tradition in the natural sciences. This tradition builds on two important assumptions. The first holds that the phenomena studied are independent of the observer, so that observations do not affect them in any significant way. In this tradition, political scientists can approach their world of study confident that there is an objective subject of study—a Real World—and that this subject will reveal itself consistently to each observer. Because the Real World exists independent of us, and because that World is stable and unchanging, it is possible to arrive at singular truths about the nature of this Real World, given the appropriate techniques.

The second common assumption holds that this Real World is patterned—that it obeys a certain logic of its own—and that these patterns are also independent of their observers. Because these patterns are a natural part of the world, this can be called a *naturalist* ontology, perspective, or worldview. While some of these patterns are patently obvious (e.g., heavy items, when dropped, fall to the ground; the sun rises in the east and sets in the west), others are concealed by the complexities of nature and life. The latter is especially true in the political world. It is the fixed existence and independent nature of these underlying patterns that allows us to generalize about the nature of social relationships. In recognizing these fundamental components of this ontology, researchers have developed methods that can search for, uncover, and eventually explain the systematic patterns that they believe exist but that are often hidden from the casual observer.

Belief in the existence of these natural and stable patterns can explain the predominance of two main sources of knowledge in political science: reason and experience (or sensory perception). These are not the only sources of knowledge available to political scientists, nor are they exclusive to the naturalist's world view. Still, rationalism and empiricism remain political scientists' most common and dominant sources of knowledge. The first of these, rational knowledge, acquires understanding about the world through strict adherence to rigorous thinking (e.g., by using syllogisms or formal logic). Because we believe the world to be stable and patterned, we can use reason and logic to uncover the underlying form, ideal, or model on which the Real World is based. While the original motivation for this way of interpreting the world was religious (that an all-powerful God created the world and imbued it with meaning), it enjoys a long and august pedigree: one that extends from Plato, through 17th-century continental European philosophy (e.g., Descartes, Spinoza, and Leibniz), and continues today in mathematical and/or rational choice approaches to studying the political world.

In the rationalist tradition, the patterned nature of the Real World can be uncovered and understood by creating consistent and logical models to show how the world really is beneath the confusing complexity of everyday life. By filtering out the particular irregularities of experience and by focusing on the underlying concepts and relationships, rationalist approaches are able to create pristine models of social behavior. Thus, for example, economists can generate models of what a true market, in equilibrium, should look and behave like (e.g., in the absence of organized interests and government regulations). Similarly, social choice theorists can show (under a number of very reasonable conditions) that no voting system is fair.

Implicitly, the authors of these models believe that they correspond to the true nature of the (underlying) Real World. To complicate matters, this Real World might differ from the world of experience. When models do not actually correspond to observable actions or patterns, supporters can argue (like Plato) that the particular experiences of everyday life (or our description of these experiences) get in the way of, or obscure, our perception of the underlying patterns and relationships. Consequently, the standard of proof in this tradition tends to be reason, not a reference to how things actually are in the world of everyday experience.

The second main source of knowledge—empiricist knowledge—is arguably the most dominant in contemporary political science. The power of this type of knowledge is derived from the belief that the most secure means of obtaining knowledge about the world is through sensory perception and the clear description of these sensory experiences. Although it may be harmful to depict the history of science in simple—dichotomous—terms, it is common and convenient to depict the empiricist tradition as a response to—or in dialogue with—the rationalist tradition. Thus, Aristotle is seen to have reacted to Plato's rationalism in beginning his treatise on *Politics* with the simple claim that "observation shows us," and British philosophers such as Francis Bacon, John Locke, and David Hume are said to be responding to the continental philosophers associated with the rationalist tradition.

Here, too, the messiness of everyday life can cloud the underlying patterns:

> The universe to the eye of the human understanding is framed like a labyrinth, presenting as it does on every side so many ambiguities of way, such deceitful resemblances of objects and signs, natures so irregular in their lines and so knotted and entangled. And then the way is still to be made by the uncertain light of the sense, sometimes shining out, sometimes clouded over, through the woods of experience and particulars; while those who offer themselves for guides are (as was said) themselves also puzzled, and increase the number of errors and wanderers. (Bacon, 1620/1863a, Preface)

Although the true patterns of a Real World lie hidden in this labyrinth of experience, empiricists believe that the use of appropriate methods and careful study can uncover them and reveal the truth. Following Karl Popper, naturalists in the empiricist tradition believe that a statement is true if it accurately corresponds to a state of affairs in the Real World.

For the world of political science, as for the world of nature (or perhaps even more so), the complexity of overlapping vocabularies and perspectives can make it very difficult to see the underlying and stable patterns that make up the Real World. To uncover these patterns, researchers must choose their approaches carefully and dissect the world systematically. Modern empiricist approaches draw from the "experimental methods" in John Stuart Mill's (1891) *A System of Logic* to develop inductive approaches that can identify the lawlike relationship in the underlying (political) world. Modern political science—in all its variants—is remarkably indebted to Mill's methods of experimentation. His method of difference, method of agreement, indirect method of difference, and method of concomitant variation have become the building blocks of modern comparative and statistical analyses: tools that are used to systematically control and compare our experiences of the political world in ways that allow us to discover the stable connections and causal regularities that are hidden beneath the surface.

Both these mainstream sources of knowledge (rationalism and empiricism) assume that the Real World not only exists but that it is naturally structured or patterned in a logical, meaningful way. Methods that rest atop both rationalist and empiricist foundations are designed to uncover singular truths about the nature of the Real World and to do so in an objective, disinterested manner. It is this pattern (or logic) that allows us to use systematic tools of induction (or methods of experimentation, to use Mill's terms) to tease out the patterns that lie just beneath the surface or to generate logical and rational models of how the world *actually is*. In this mainstream (naturalist) tradition, political scientists can transcend their personal, temporal, and cultural perspectives, and beliefs and prejudices, to arrive at valid knowledge and objective truths.

Most contemporary political scientists have inherited approaches that have grown out of this tradition: They have found utility and meaning in methods that have been carefully adapted in ways that make them more suitable for studying the political world. Indeed, scholars in this tradition subscribe to a strong (if implicit) hierarchy of methods for uncovering truths about the world, and they do so by combining elements from both the rationalist and empiricist traditions.

At the top of this hierarchy lies the experimental method. This method is prized because it allows researchers to manipulate the relevant variables, in a context that controls for all other sources of influence, in order to secure firm knowledge about posited relationships. This method works because we assume the world to be made up of discrete and independent parts (not unlike those in a clock), which can be mixed and matched in different ways to reveal (and test) causal relationships. Indeed, this capacity to control and compare components of the real world is a hallmark of all good science.

When experimentation is not a realistic alternative—and many areas of political life do not lend themselves to experimental design—naturalists tend toward statistical approaches. Here too, manipulation, comparison, and control are central to uncovering real-world patterns. But instead of manipulating the contexts in which these variables lie, the statistician's computer manipulates the data in ways that mimic experimental control. Because statistics do not involve the physical manipulation of data, it is a method that lends itself to the study of social phenomena, where the tendency is to study events that have already occurred.

The naturalist's third best alternative is small-*N* comparative approaches. These are used when one is interested in questions that have too few observations to run reliable statistical enquiries. As both statistical and small-*N* comparative methods use experimental designs, they share much in common. The most important difference is that the comparativist selects the cases of study in a strategic way, so as to maximize the variance of the independent variables and to minimize the influence of other (control) variables. In other words, unlike statistical and experimental designs, case selection in small-*N* comparative projects is anything but random. Because of this, researchers need to rely on theory to avoid major problems associated with this method and its control strategy (such as overdetermination and a sampling bias).

Finally, case studies constitute the bottom rung of this methods hierarchy. This method is used only when researchers are faced with a paucity of data or relevant comparisons. Case studies are histories with a point: They gather information on cases that are interesting or relevant in light of a larger theoretical concern or a specific research design. By collecting more data on such cases, researchers might eventually bump the study up to a higher rung on the hierarchy. On their own, however, case studies have little value, in that studying a single case can provide little grounds for generalization about the patterns that motivate the research.

These methods work remarkably well when used in the right contexts and when the limitations of each are recognized and respected. These contexts draw on a naturalist ontology that sees the world as something independent of the observer, patterned in nature, and accessible by methods that use observation, control, and comparison. The methods depend on a world that can be broken down into its component parts (variables), which are themselves independent and autonomous. This allows the researcher to cut up a political phenomenon, manipulate its component parts, and use this structured manipulation (of either the world or the data) to interrogate the phenomenon in a way that will reveal its logic or pattern. However, when these methods are used in a methodologically naïve manner—when they are used in contexts that do not correspond to the assumptions that gird them—the results can be misleading, biased, and wrong.

A Few Examples

There is little point in humiliating those who have ignored the ontological assumptions that underlie their methods. This problem is common enough for the shame to be shared by many. Still, this point can be made with reference to a couple of recent (positive) examples of colleagues who have reflected on (and responded to) the shortcomings of a particularly popular method. Although both these examples are critical of cross-national regression analyses, it is not meant to suggest that this problem is confined to statistical studies. These examples are used because they highlight the sort of difficulties that can be found, even among the most refined, explicit, and sophisticated of contemporary methods. A fundamental assumption of most comparative research is that a representative sample of independent observations, drawn from the population of interest, is used. The independence of observations is important as it affects the effective number of observations (e.g., if there are four people in a household and they generate an opinion together, the number of independent opinions is one not four). Samples based on observations that are not independent will generate findings that misestimate the amount of variability in the population, producing biased sample statistics and erroneous conclusions. This assumption of (observational) independence meshes well with the ontological assumptions girding many of the fields of research that use regression techniques. When applied on international samples, this assumption is also in tune with the logic of the Westphalian state system (at least an idealized view of that system), where states are understood to be autonomous and independent. But in a world that is increasingly characterized by international integration and diffusion, this assumption is increasingly tenuous.

This problem is nothing new, and it is usually traced back to the critical comments of Sir Francis Galton, who—late in the 19th century—questioned the results of a paper that used data from a cross-cultural sample to test the relationship between marriage laws and descent patterns in tribal cultures. “Galton’s problem” was born when he noted that the observed correlation might have been a result of contacts between the cultures in the sample (and not the result of a test of truly independent cases). While anthropologists have been struggling with Galton’s problem for some time, political scientists have been remarkably slow to realize (and address) the problem. In recent years, however, there has been an explosion of activity on this front, as many statisticians attempt to deal with the effects of globalization. Detlef Jahn (2006) provides political scientists with a good introduction to the nature of the problem and suggests one way to try to deal with it. In an honest attempt to align method and ontology, Jahn proposes a simple modification of ordinary least-squares (OLS) regression techniques that allow him to measure the strength of the diffusion associated with globalization.

This is one way to resolve the growing tension between a changing world and the most popular method for its study. Jahn provides us with a "patch" to an approach that is not particularly well suited for dealing with the changing world of international politics. These sorts of patches work if they are few in number and do not overwhelm the underlying approach (the result, then, would be like trying to run Windows XP on an old 386 processor). In the long run, it is likely that an approach that is better designed, from the bottom up, will need to be developed to deal with this new context. In the meantime, however, Jahn gets us thinking.

A second example comes from an article by Peter Hall (2008). Like Jahn, Hall is critical of mainstream regression techniques but for different reasons. His concern is not so much in the changing nature of the things researchers study but on the different kinds of things they study. In particular, Hall is interested in how researchers secure causal explanations of political phenomena. More to the point, Hall is critical of the (common) assumption that large-*N* statistical approaches provide the best grounds for causal inference. After reviewing three different modes of explanation (historically specific, multivariate, and theory oriented) Hall shows how small-*N* comparative designs can be valuable for causal inference. In particular, an intensive examination of the causal chain provides researchers with a new and different basis for causal inference—one that is especially well suited for assessing the sort of complex causal theories that are so prominent in the social sciences. When they use *systematic process analyses* (Hall's term), small-*N* comparisons can be especially useful for developing theory-oriented explanations.

Hall's reasoning rests on ontological foundations. In addition to traditional arguments (such as the need to be aware of the state of the literature), Hall notes how the researcher's methodology should be conditioned on the state of the world, as he or she perceives it (Hall, 2008; see also Hall, 2003). Because standard regression models build on assumptions that are unlikely to hold in situations where researchers are trying to develop theory-oriented explanations, he questions the (common) claim that a single (statistical) method can adequately assess the validity of causal inferences in the social sciences. While the underlying assumptions of standard regression analyses

> were usefully employed to assess the conditions conducive to securing stable democracy, when those conditions were thought to include basic socioeconomic factors, such as the level of economic development, related levels of literacy and the correlates of "modernisation" . . . when theorists began to see stable democracy as the product of an intricate strategic interaction among reforms, extremists and defenders of the old regime, statistical methods were no longer appropriate for assessing the causal chain. (Hall, 2008, p. 307)

Both these examples show how greater awareness of ontological assumptions should be important for empirically minded political scientists. In the first example, this sort of awareness was used to tweak an existing approach in a way that makes more sense in a changed (and changing) world of study. In the second example, explicit reflection on the different types of causal arguments was used to argue for the inclusion of small-*N* methods, more suitable for some of the tasks at hand. Both examples illustrate the importance of embracing different methods and thinking carefully about how our methods engage the world we study.

Competing Worldviews

As with just about everything else in political science, there is more than one way to look at the political world. Indeed, there are many who are unwilling to commit to a single Real World of political study, and this has important consequences for the way that methods (even familiar methods) are used. While the focus of this entry is based on mainstream, naturalist approaches to political science, it can be useful to glance at an alternative ontological perspective to show how standard methods play different roles when used in different contexts.

This alternative ontological perspective can be called *constructivism*. Its followers question the existence of a single world of study, independent of its examiner and capable of revealing objective and singular truths. Instead, constructivists believe that truth lies in the eyes of the observer and in the constellation of power and forces that support that truth. As our observations and descriptions of events cannot be shielded from the biases that surround us, constructivists do not believe that it is possible to secure an absolute truth that corresponds to a single Real World. Their objectives are more modest: They point to the ways in which our contexts (and those of our subject matter) frame the way we come to develop an understanding.

This constructivist tradition is often marginalized by mainstream political science for being unreliable and even serendipitous. Some think mainstream political science does so at its own peril. This is because constructivism shares a number of ontological and epistemological features with the naturalist tradition, and these similarities allow each tradition to borrow from the other. In some respects, the two perspectives can complement one another, and the lessons learned from one tradition can be amended, or modified, to be applied in contexts more familiar to the other. Indeed, this useful collaboration is evident in the growing movement for scientific realism. Most important, constructivists and naturalists share a willingness to see the world in terms of interpretable patterns, and both traditions rely heavily on rationalist and empiricist sources of knowledge for understanding these patterns.

How these perspectives differ is in their view of the source of these patterns and the ends to which they use rationalist and empiricist sources of knowledge. For constructivists, the patterns of interest are not firmly rooted in nature, but they are the products of our making: Each of us sees and experiences different things, and what we see (or experience) is determined by a complicated mix of social and contextual influences and/or presuppositions.

The modern intellectual roots of this tradition can be traced to Immanuel Kant, who was struggling with the implication of David Hume's difficulty in securing firm knowledge about causality. Hume's approach to causation rested on a rather simple theory of sensory perception: The human mind (straightforwardly, even mechanistically) absorbs impressions through the senses. Kant seems to have appreciated Hume's general theory of sensory perception, but he did not think that the human mind should be understood as an empty vessel into which sensory perceptions were merely dumped. For Kant, our senses only bring perceptions to the doorstep of the mind. It was then up to the mind to reorganize, characterize, and then store these perceptions for later use. To perform this task, Kant believed that the human mind was prewired with basic preconditioning concepts, or forms of understanding. Thus, in thinking about how we experience and store sense perceptions, Kant shifted the ontological terrain of patterned (lawlike) behavior from nature to the human mind. In other words, Kant believed that the patterns we see in nature are the result of our mental manipulation of sensory perceptions—not a characteristic of nature itself.

Kant did not suggest that the social world was ontologically different from the natural world. Neither did he question the existence of a Real World. Rather, Kant argued that we cannot know anything about the Real World (*noumena*, in his vocabulary). All we can know is that our perceptions (*phenoumena*) of the Real World are somehow related to it—and that the nature of this relationship is complex and ambiguous. The result of this complex and ambiguous process is that scientists (both natural and social) find themselves facing what appear to be different, even competing, worlds.

There are several other reasons to question the naturalist ontology. Many of these responses emphasize the unique qualities of human agency to argue that the nature of human relations must be different from those that we find in the natural world. "We are cultural beings, endowed with the capacity and the will to take a deliberate attitude toward the world and to lend it significance" (Weber, 1949, p. 81). Constructivists take observations such as these to show how human agency, itself, can create objects that have a different ontological status than those found in the natural world: where the very existence of social facts (e.g., money, property rights, and sovereignty) depends on human agreement and institutions. Given this different ontological context, constructivists question the utility of employing methods to search for naturally occurring (independent) patterns of behavior.

Because constructivists embrace a world that is ontologically diverse and complex, they tend to draw on more and different types of evidence and proof. While constructivists draw heavily on rationalist and empiricist sources of knowledge, they are often willing to employ other, more radical, sources of knowledge as well (including, e.g., empathy, revelation, and myths). While these alternative sources of knowledge may be less reliable, they are firmly anchored in human experience. For this reason, they can provide important keys for unlocking the (constructed) patterns researchers study. As constructivists are not searching for objective and verifiable patterns of phenomena existing naturally in the social world, they do not subscribe to the naturalist's hierarchy of methods. After all, this tradition doesn't believe that truth is just "out there"; knowledge about the social world is always knowledge in context; it is socially situated and it has social consequences. While it is difficult to argue that constructivists subscribe to an explicit methods' hierarchy of their own, they clearly prioritize methods that protect and nurture the contexts that can explain the source of the patterns they see. As these patterns are understood to be socially (or individually) constructed, their methods aim to uncover the motivations and presuppositions that generate the underlying patterns.

For this reason, constructivists have a soft spot for narrative approaches, as these provide researchers with a nearness to the data and context that is necessary and desirable in order to gain insight. This emphasis on narration and the importance of context and contingency is extended to the way in which constructivists employ comparisons—as a tool for developing associations that can help establish meaning. While constructivists seldom use statistical and experimental approaches, there is nothing inherent to either of these methods that need alienate constructivist scholars. With a little thought and imagination (two traits usually embraced by constructivists), statistical and experimental approaches could be designed and used in ways that can exploit contextual familiarity and appeal to constructivists. In short, constructivists tend to use the same basic methods, but they do so in different ways, toward different objectives.

In a nutshell, naturalists tend to exhibit little regard for the surrounding context. Their methods are designed to cut into contexts in order to gain access to the component parts. These are then manipulated to control for, and capture, relevant variation. Because the patterns they seek are assumed to exist independently of the component agents and their observers, preserving the interpretive context is not a high priority for naturalists. Constructivists, on the other hand, want to use comparisons to better understand the nature or source of these patterns. As the patterns they see are a function of the agents they study, or the observers themselves, familiar methods are often turned inside out to look critically at how the contexts themselves provide meaning to the patterns under study.

As in the naturalist tradition, constructivists can draw from an impressive intellectual legacy, which stretches back to antiquity. More surprising, perhaps, is that this perspective finds support among some rather unlikely allies. For example, this entry has already noted how indebted mainstream (naturalist) political science is to the experimental methods developed by J. S. Mill. What was not mentioned previously was that Mill himself (e.g., Mill, 2002) was skeptical of their application to the world of politics, and his skepticism rested on ontological foundations.

This sort of skepticism has been voiced by many others. Indeed, these shaky ontological foundations may be the reason why political science knowledge doesn't measure up (in terms of, say, the ability to generalize and/or predict) to the knowledge produced in the other branches of (natural) science. While such failures might be explained by other factors (e.g., the youth of the discipline, the complexity of human action), there are formidable arguments—authored by reputable thinkers such as Wilhelm Dilthey, John Stuart Mill, Martin Heidegger, Jürgen Habermas, and Charles Taylor—for believing that the shortcomings of this discipline are the result of relying on methods that rest on faulty ontological assumptions.

These sorts of challenges are important, if only because they force political scientists to consider ontological alternatives. The very existence, and growing popularity, of these alternatives should encourage political scientists to consider the degree to which their approaches and methods rely on assumptions about the nature of the real world. After all, if the political world is not really as political scientists assume—if its patterns are not constant, universal, and independent of us—then the workhorse methods they have developed will lose much of their pull.

Overcoming Bias

The previous two sections argued for the importance of thinking carefully about the nature of the political world researchers study and how that nature affects the way they can and should study political phenomena. This section sets aside these larger ontological issues and narrows the focus to the most common approaches in contemporary political science: approaches that assume a naturalist ontology and rely on empiricist sources of knowledge. But even here, in this familiar ontological and epistemological context, political scientists face a number of important obstacles that can stand in the way of securing objective truths about the world they study. Overcoming these obstacles is another role to consider when choosing a particular method or methods.

While there are many ways to think about the practice of securing unbiased and true knowledge from the Real World, most of them draw from a common and distinguished intellectual legacy. This section highlights the contribution of a foundational thinker in that legacy: Francis Bacon. Bacon's approach to studying the political world is similar to that of contemporary political science: It begins by assuming that there is a real world out there and that it is possible to obtain true knowledge about that world. Also, Bacon was not particularly troubled by the possibility that our senses might deceive us (or any of the other arguments advanced by those who are skeptical of our ability to secure truth). Rather, Bacon was concerned with unmasking the underlying sources of misconception, irrationality, and error that bar us from reaching the truth. He seems to have believed that any impediment to knowledge could be overcome by the careful and considered means of dealing with the subtleties of nature:

> The mind of man is far from the nature of a clear and equal glass, wherein the beams of things should reflect according to their true incidence; nay, it is rather like an enchanted glass, full of superstition and imposture, if it not be delivered and reduced [corrected]. (Bacon, 1605, Book 2, chap 14, p. 9)

To deliver and reduce this enchanted glass, Bacon provided us with a typology for thinking about the most common hindrances to seeing truth. In *Novum Organum* (1620/1863b), Bacon introduces the idols of the mind that had hitherto blocked the progress of mankind's knowledge of nature. For Bacon, these idols are like deceitful pictures—they are images that stand in the way of our grasping the truth, even though they may be grounded in human reason. Together, these idols capture the range of mental, psychological, and social dispositions that Bacon believed were responsible for distortion and error. The first of these are the *idols of the tribe*. These idols are intellectual errors that result from our common (i.e., human) weaknesses. For example, Bacon tells us that we tend to assume more order and regularity in the world than is actually the case; we have a soft spot for common sense—that is, we rely on limited assumptions without trying to verify them as established truths; we tenaciously hold on to our beliefs by means of continual rationalization, even in the face of countervailing evidence; we are attracted to positive examples more than to negative examples; we pursue wishful thinking (i.e., we believe what we want to believe); and we have a tendency to overgeneralize and to believe in ultimate causes.

In short, Bacon's idols of the tribe refer to our tendency to project our own (human) patterns of instinctive thinking onto nature. This tendency has not died with Bacon, nor has it been permanently corrected by subsequent methodological developments. Indeed, these idols can be seen at the forefront of all the social sciences. For example, the idols of the tribe might be used to explain why political scientists search for (and find!) the democratic peace, why economists find it fruitful to borrow the concept and mathematics of "energy" from 19th-century physics and apply it to the concept of utility, and why social science journals do not appear to be interested in publishing negative findings (and why it has been so difficult to launch a journal of negative results in the social sciences). More to the point, these idols may explain our willingness to assume that nature is constant, independent, and patterned, even in the absence of an almighty God that was once said to have made it so. Bacon himself would have political scientists continually question the underlying ontological assumptions of their discipline—not because those assumptions are unique to their discipline but because those assumptions might reflect a deeper, human need for a world that is constant and patterned, as if designed by a god.

This entry is not suggesting that the democratic peace, neoclassical economics, or all of the positive examples in professional journals are either false or illusive. (Nor does this entry intend to argue for—or against—the existence of God.) The point, like Bacon's, is more modest: Researchers need to be extra diligent and skeptical when they discover anthropocentric patterns and arrive at convenient truths.

The second category of Bacon's idols concerns those of the cave. *Idols of the cave* represent the peculiarities of each individual's temperament and limitations. These individual biases can include, for example, tendencies toward insularity, conservatism, or novelty; acceptance of authoritative propositions; or a willingness to bypass intellectually difficult positions. While the idols of the tribe are derived from the human condition, the idols of the cave reflect one's own individual biases. These biases are particularly difficult to shed, as they are derived from (and forged in) individual experience. But they are biases, nonetheless. Consider how practitioners of political science often become beholden to a particular method. As methods become more specialized and technical, this sort of method's monogamy is increasingly common. This specialization is, in itself, a welcome development, as a division of specialized labor can boost productivity and the quality of its research. In doing so, however, political scientists reverse the way in which scientists have traditionally understood the relationship between the world, the methods, and the understanding. Rather than let the nature of the question determine the proper method and means of understanding, we see the world as we would have it from the comfort of our cave.

This method's monogamy can have serious consequences for the way we come to interpret and understand the world. Political scientists may remember their own training in regression techniques and how exciting (and challenging) it was to interpret the world anew in terms of dependent and independent variables. This new perspective can be liberating, but it is also blinkered: "Someone with a new hammer thinks the world as a nail."As with the idols of the tribe, Bacon teaches us that it is possible to overcome idols of the cave by being aware of our individual biases and confronting them directly in a critical, balanced, and clear way. It is for this reason (rather than a need for a clear demarcation principle) that researchers should strive to falsify their findings.

The next two types of idols are not so easily corrected. Bacon's third category, the *idols of the marketplace*, is derived from the shortcomings of language and intercourse, as words often betray their own purpose and obscure the very thoughts they are designed to express. At the most general level, Bacon's idols of the marketplace contain two types of misunderstandings. First, we often give confusing, even ill-defined, names to things that do exist. After all, the world is full of meaningless words, words with double meanings, personalized meanings, jargon, and so on. Second, names are sometimes given to unreal things (e.g., Fortune, Prime Mover). While the first type of errors is difficult to eliminate, as they are deep-seated and quite complicated, Bacon believed that the latter errors can be easily thrown out along with the theories that inform them.

Most social scientists are aware of the slippery and inaccurate ways in which language and concepts can be used in the social sciences. Consider, for example, the different ways in which "significance" is often used in statistical papers or the varied and overlapping ways that political scientists refer to democracy, political parties, globalization, or even justice. As Émile Durkheim noted,

> In the present state of knowledge, we cannot be certain of the exact nature of the state, of sovereignty, political liberty, democracy, socialism, communism, etc. Our methods should, then, require our avoidance of all use of these concepts so long as they have not been scientifically established. And yet the words which express them recur constantly in the discussions of sociologists [and political scientists]. They are freely employed with great assurance, as though they correspond to things well known and precisely defined, where as they awake in us nothing but confused ideas, a tangle of impressions, prejudices and emotions. (Durkheim, 1895/1964, pp. 65–66)

The fourth and final of Bacon's categories contains the *idols of the theater*. These are idols derived from things learned in the past, on the basis of poorly conceived experiments and superstitions. This makes them particularly difficult to avoid or overcome, as false learnings—when cultivated—gain wide influence and become unquestioned by their followers. It is because of this that

Bacon dedicated most energy to deal with these idols, criticizing the habits of earlier philosophers and their approaches.

Presently, the political world (and its most influential actors) has begun to recognize the idolatry of a once dominant model of political and economic exchange. The idea of a self-regulating, equilibrating marketplace—one that could usefully and peacefully coexist within democratic polities—had become ingrained in the minds of men and is supported by the sort of poorly conceived experiments and superstitions to which Bacon referred. As the curtain appears to be falling on this idol of the theater, political scientists will have to rely on new understandings of the political world—understandings that are more aware of their debt to the "various dogmas of philosophy" and the "wrong laws of demonstration."

For Bacon, these idols did not challenge his underlying (naturalist) ontology. Rather, he sought to undermine the authority of accepted sources of knowledge; he wanted to point out the pitfalls that lie along our path to more useful knowledge, as only then do we have a chance of avoiding them. Bacon believed that the mind of man is like a circus mirror; it reflects what is going on in the world, but it distorts it in the process. As a result, our mind produces false images or idols. While the mind will always create these false images, Bacon believed that it was possible to compensate for the distortion by building a correcting lens. Broad-based methodological reflection provides that lens.

Conclusion

Bacon's *Novum Organum* is divided into two books. The first book introduces the idols of the human mind, to show the sort of obstacles, biases, prejudices, and limitations of perspective that can stand in the way of our quest for knowledge. These idols warn of the dangers inherent to both inductive and deductive approaches.

The second book provides the sort of lens that he believed was necessary to correct for the maligned images. In doing so, he did not argue that all the idols could be overcome. Indeed, the latter two types of idols (those of the marketplace and theater) could not be eradicated, as they denote social and ideological biases that are imposed from the outside. For these idols, we can only be aware of their existence and hope that this awareness will protect us from the insidious effects they have on our minds.

The first two idols, by contrast, "inhere in the nature of the intellect." For this reason, Bacon believed that these idols of the tribe and the cave could be overcome by confronting them in a critical, balanced, and clear way. It is in this light that methodological reflection is so important for good science—whatever its ontological foundations. Bacon's solution to these problems was to develop a single method, called induction, which could correct these many distortions. This name for his method has fooled many into seeing Bacon as some sort of poster boy for atheoretical empiricist approaches. It is hoped that by reviewing his idols of the mind, readers will see how critical Bacon was of the idea that simple observation alone could convey accurate knowledge of the Real World. Bacon's (1620/1863b) approach combined rational and empiricist elements in a middle-course approach:

> Those who have handled sciences have been either men of experiment or men of dogmas. The men of experiment are like the ant, they only collect and use; the reasoners resemble spiders, who make cobwebs out of their own substance. But the bee takes a middle course: it gathers its material from the flowers of the garden and of the field, but transforms and digests it by a power of its own. Not unlike this is the true business of philosophy; for it neither relies solely or chiefly on the powers of the mind, nor does it take the matter which it gathers from natural history and mechanical experiments and lay it up in the memory whole, as it finds it, but lays it up in the understanding altered and digested. Therefore from a closer and purer league between these two faculties, the experimental and the rational (such as has never yet been made), much may be hoped. (Book 1, p. 95/xcv)

In the same way that Bacon counsels us to borrow from the ways of the ant *and* those of the spider, good methodologists need to be aware of—and draw from—competing worldviews, sources of knowledge, and methods. Choosing the right method requires critical reflection on how a given method can interrogate the world under study and generate the type of knowledge we hope to obtain.

Jonathon W. Moses
Norwegian University of Science & Technology
Trondheim, Norway

See also Causality; Constructivism; Quantitative Methods, Basic Assumptions; Quantitative Versus Qualitative Methods

Further Readings

Bacon, F. (1605). *The advancement of learning*. New York: Dent & Sons. Retrieved from http://www.uoregon.edu/~rbear/adv2.htm

Bacon, F. (1863a). *The great instauration*. In J. Spedding, R. L. Ellis, & D. D. Heath (Trans.), *The works* (Vol. 8). Boston: Taggard and Thompson. Retrieved November 27, 2010, from http://www.constitution.org/bacon/instauration.htm (Original work published 1620)

Bacon, F. (1863b). *Novum Organum*. In J. Spedding, R. L. Ellis, & D. D. Heath (Trans.), *The works* (Vol. 8). Boston: Taggard and Thompson. Retrieved November 27, 2010, from http://www.constitution.org/bacon/nov_org.htm (Original work published 1620)

Durkheim, É. (1964). *The rules of sociological method and selected texts on sociology and its method* (G. E. G. Catlin, Ed.; S. A. Solovay & J. H. Mueller, Trans.). New York: Free Press. (Original work published 1895)

Hall, P. A. (2003). Aligning ontology and methodology in comparative research. In J. Mahoney & D. Rueschemeyer (Eds.), *Comparative historical analysis in the social sciences* (pp. 373–406). New York: Cambridge University Press.

Hall, P. A. (2008). Systematic process analysis: When and how to use it? *European Political Studies, 7,* 304–317.

Jahn, D. (2006). Globalization as "Galton's problem": The missing link in the analysis of diffusion patterns in welfare state development. *International Organization, 60,* 401–431.

Mill, J. S. (2002). *A system of logic*. Honolulu, HI: University Press of the Pacific. (Original work published 1891)

Moses, J. W., & Knutsen, T. L. (2007). *Ways of knowing: Competing methodologies in social and political research*. Hampshire, UK: Palgrave.

Shapiro, I. (2005). *The flight from reality in the human sciences*. Princeton, NJ: Princeton University Press.

Topper, K. (2005). *The disorder of political inquiry*. Cambridge, MA: Harvard University Press.

Weber, M. (1949). Objectivity in social science and social policy. In E. A. Shils & H. A. Finch (Ed. & Trans.), *The methodology of the social sciences* (pp. 50–112). Glencoe, IL: Free Press. (Original work published 1904)

EQUALITY

Throughout human history, equality as an ideal has evoked powerful emotions. It has fueled revolutions and social movements, and people have died and have killed for it. In addition, it has been the subject of formal analysis that has sought to analyze, compare, and measure inequalities in goods, liberties, and opportunities. This entry examines the concept of equality, beginning with the two fundamental questions: "why equality?" and "equality of what?" The second issue concerns the varied domains in which equality may be sought. This distinguishes the various social ethical conceptions, but it seems a priori that such a moral question can be answered only if the former question is: why equality in the first place? There turns out to be various types of reasons for equality, and in particular the trivial equality as generality or universality, the basic equality as logic or rationality, and the socially essential equality as the absence of subjection and domination (hence as this freedom). The discussion of equality from rationality introduces an analysis in terms of rule equality.

The entry then explores alternative conceptualizations of distributive justice, which can be considered with respect to both their substance (the goods to be distributed) and their structure (the principles according to which distribution should occur). The concept of equality of liberty is then explored with respect to the equality of basic rights; of real liberty (including means for the possibility of free action); and of opportunity. The relationship between equality and responsibility, the concept of equal hypothetical liberty, and the notion of comparative egalitarianism are also considered.

The concept of equality is an important concern for political science, and the entry considers the relationship between democracy and political equality. For philosophers, the concept of impartiality has been a central one, and the entry discusses its elaboration through the theory of the original position. The work of social scientists on the measurement and comparison of inequality is also noted. Finally, equality in social relations can be seen as grounded in reciprocity—the fundamental tendency of human beings to treat and relate to others as they treat and relate to you.

The Problem of Equality

Equality as the First Virtue of Society

Human history is intrinsically linked to inequality, in such forms as slavery, racism, apartheid, sexism, class and caste structures, and other forms of domination and discrimination. A vast array of facts, emotions, and reasons establishes the overwhelming importance of both the question of equality and of its necessary conceptual clarification. Equality may sometimes be so bad that only one thing can be worse: its absence. "Inequality is the source of all evil" is Jean-Jacques Rousseau's (1755) clear-cut conclusion of a nevertheless elaborate investigation. Aristotle and John Rawls see justice as actual or ideal equality and find it to be the first virtue of society. Indeed, "Justice is equality, as everybody thinks it is, quite apart from other considerations" is Aristotle's teaching to the king's son in *Nicomachean Ethics*. Social ethical equality, the topic here, is almost consubstantial with the concepts of justice in the same field (social justice, distributive justice, compensatory justice, restorative justice, commutative justice, diorthic justice, etc.), but this entry first considers the issues from the equality angle.

Equality of What?

Of course, equality can a priori be of many things, with often opposite actual consequences. Equality may be of various types, including

- equality of income or goods;
- equality of basic rights or liberties;
- equal right to the product of one's labor or capacities;
- equality in psychological welfare (happiness, satisfaction);
- equality in "real" freedom of choice;
- equality of opportunity;
- the political equalities;
- relational equality;
- equality in dignity;
- equality of consideration;
- a society of equals; and
- ontological equality.

Equality is commonly thought to mean equality in incomes or goods. It can also be in liberty, however. Historically, in fact, the first and main demand for general equality was equality in rights and notably in basic rights, which are essentially liberties ("Men are free and equal in rights" is the opening statement of the 1789 Declaration of the Rights of Man and of the Citizen). This freedom from forceful interference has been seen as forbidding income redistribution, thus as meaning equal full self-ownership (i.e., each person is entitled to the effects and products of her own capacities to work and earn, and to enjoy) and hence implying a precise opposite of income equality. Equality may be not in goods but in the (psychological) welfare or "happiness" people derive from them thanks to their capacities to enjoy. On the contrary, it can be in the (other) resources given to society, and therefore in the "real liberty" of using them, thus complementing the "formal liberty" provided by the basic rights. If these resources are attached to the individuals, as their earning capacities or social conditions are, transfers or specific policies achieve this equalization. This can give various equalities of opportunity. Equality can also be in the variety of social relations, processes, statuses, situations, or conditions. Of particular importance is political equality, equality in political power and civic duty, and its manifestation in democracy. Finally, one kind of equality is particularly fundamental in the ethics of modernity: that of the basic moral worth of humans as such, with the attached respect, consideration, dignity, and social and material consequences. This "ontological equality" refers to our common humanity, which should be respected in all its instances (basic moral equality). In Immanuel Kant's words, all humans are equal in the kingdom of ends, and no one should consider any other as a means only. Equality can also appear in different types of rules that permit to determine individual situations. It is, for instance, an equality of weights in utilitarianism or in the highest social income (highest sums of individuals' utilities or incomes). Equality is also sometimes *rule-equality* (or *functional equality*), that is, the items of individuals are derived from their specific given or chosen characteristics by the same rule or function. As is noted below, this is the very structure of rationality in the sense of providing a reason, with important consequences.

Equality and Modernity: Formal and Real Equalities

The equalities considered here are results of choices by society, often by institutions but sometimes by individuals. In almost all societies there are peer groups with some values of equality between their members, and, often, equalities of certain types with larger extensions. However, we are also particularly interested in equality in the ethics of modernity. The logical analysis of equality will apply to all cases. The ethics of modernity is characterized by the acceptance or demand, by large majorities of populations, of certain equalities for large populations, universally for some equalities. These ideal values are, first, moral basic worth, classical basic rights, and some sort of democracy. Respect, and basic rights when the distribution of resources is given, are nonrival (that is, one individual's benefit from an item does not prevent or impair similar benefits for others) and therefore the demands may simply be that each person should have them, which implies their equality. In contrast with these consensual values of the ethic of modernity, this ethic is deeply divided with regard to the distribution of goods—the economic values. The polar positions are, on the one hand, a divided family of "egalitarians" who favor equality in incomes, goods, resources, or welfare and, on the other hand, "classical liberals" who advocate self-ownership of all personal capacities (to earn and enjoy)—and hence, by the way, equal self-ownership for all. This issue and the resulting structure of the optimum distributions are discussed below. Note that since (*prima facie*) equal treatment of equals in the relevant characteristics turns out to be a logically necessary property of a determinate social choice with minimal rationality (cf. below), equality appears in two different ways in social choice: (1) as this rationally necessary property of all social ethics that applies to the particular equalizand and scope of this ethics whatever they are—it can in particular be an equal freedom, for instance; and (2) as the particular values of the noted family of distributional "egalitarians" (in goods, incomes, resources, or welfare).

The object of equality has various possible structural properties, discussed below. For now, simply note that equality can be between individuals but also between groups or institutions variously defined (with, possibly, the problem of relating the situation of the group to that of its members). For simplicity in presentation, this entry uses expressions of equality between persons or individuals only.

Why Equality?

Equality raises two classical questions: "of what?" (including between whom and in what circumstances) and "why?" The operational question is "of what?" However, it seems that it can be answered if and only if we first have the answer to the other, apparently deeper question, "why?" The issue is more subtle, however. Consider, for instance, very common claims such as that all humans should equally have, for instance, anything such as "the basic rights," or "at least the food needed for survival." This "of what," whatever its own reason, constitutes the reason for this equality, it explains it. Then, the answer to "equality of what" entails the answer to "why equality?" In this case, the mention of equality is in fact redundant. Yet it is often emphasized for reasons noted below. In another example, the very commonly given reason "I divide this cake equally because I see no reason to divide it otherwise" has a puzzling logic analyzed shortly. In other cases, equality and its reason or value are just two different names for the same thing, as with the most important equality as nondomination or nonsubjection (the two faces of the same relational coin).

When reasons for equality are considered, the striking fact is that there is not one reason or motive for equality but many of them, of very different and often unrelated kinds. The two most important types of reason for equality are of totally different natures. One is *equality as logic or rationality*. It concerns the reason for "equal treatment of equals," the logic of justification, the property of "permutability," and the meanings of justifying equality by the absence of a sufficient reason for inequality. The other type is social. It is *equality as nonsubjection and nondomination*, a protective or negative relational equality, justified by this type of liberty and of dignity, and extending to the general properties of relations between equals. Equality as rationality can apply to all issues—economic, social, political.

Logic—if one dare say—provides also another reason for equality that is trivial from its viewpoint, a tautology, and is nevertheless often repeated, sometimes with great emphasis and a great importance attached to it. This is *equality as generality or universality*, meaning that each member of a given group has or should have some given property of any nature. This is extended into a comparison: each member has, all members have, all members equally have. This property then is general to the members of the group. It is "universal" in this group, but the term "universal" is often reserved for cases in which the group is all mankind. Logically, this mention of equality is redundant. Its presence may have two reasons aiming at reinforcing the claim or value. One is to draw attention on the fact that, in the present or past states, some members only have or had the property. Another may be to appeal to other reasons for equality, namely comparative fairness based on the logical reasons mobilized by the emphasis that the persons in question have the same relevant characteristics.

Comparative equality results from the comparison of persons' endowments of the items relevant in nature and in measure (e.g., perhaps the appropriate concepts relative to some specific characteristic of the person). Equality then results from sentiments of relative fairness, and it prevents the various social sentiments that may be aroused by inequality (such as envy or sentiments of injustice, unfairness, inferiority, or superiority). This fairness, however, is based on the notion that the persons have the same relevant characteristics (no one deserves, needs, or is entitled to or accountable for more than the other) and on the logical reasons presented in the section dealing with rationality as a reason for equality. Relatedly, the principle of "equity-no-envy" saying that each person prefers to have her own allocation or situation rather than that of any other person holds a central place in equality analyses.

Equality, therefore, is essentially a derived value. It derives from direct (end-) values by implications that are varied and opposite in type and direction. In the various cases, it is a condition, a cause, or a consequence. For nondomination, equality is factually identical with it and hence morally a consequence of it. Directly comparative approval of equality results from some sentiment of propriety perhaps supported by the justification from rationality. However, it is not sure that equality is or can be valued as an end in itself, directly, although it may look like this in some egalitarian judgments that appear as gut feelings or flashes of moral intuition, previous to considered analysis (the opposite of the search for a good reason). This may concern, in particular, the basic worth of humans, relational equality in itself (relation between equals), comparative fairness, the absence or impossibility of a reason for inequality, and the pure quasi-aesthetic value of balance and symmetry.

When the relevant equality is impossible or costly on other grounds, some reasons for it or judgments favoring it can extend to preferring lower corresponding inequalities. This extends considerably the complexity of the problem and constitutes a vast field of studies. When what is wrong with inequality is that people who have the least have too little, and if another situation can improve their situation sufficiently without costs in the other people's endowments of this item or otherwise that would make the overall situation worse, the solution may be to maximize the lowest endowments or "maximin" ("practical justice" for interpersonally comparable ordinal utilities in Serge Kolm [1971], the "difference principle" for an index of "primary goods" in Rawls [1971], or Derek Parfit's [1995] "prioritarianism").

Finally, some equalities induce, entail, or require others. This can result from the existence of strictly complementary goods. For instance, enjoying some right or liberty may require some condition such as the access to some amount of some good. But the most famous and classical example is Arthur

Pigou's derivation of equal income from the utilitarian highest sum—hence with equal weights—of identical concave individual utility functions. A more elaborate similar property is the basis of the present-day welfarist theory of measures of inequality.

The essential question of the relations between equality and liberty will be split in two: *equality as liberty*, the historically most important defensive relational equality of nonsubjection and nondomination, and *equality of liberty*, including the basic rights and the various cases of equality of freedom of choice and of opportunity.

Equality as Liberty: Nondomination and Nonsubjection

Today, equality is commonly considered as opposed to liberty. This usually refers to inequalities in income and wealth resulting from free exchange, and to interferences by public redistributions tending to reduce these inequalities. It sometimes also refers more philosophically to freedom permitting the manifestation of differences in preferences in a diversity seen as an inequality. However, liberty and equality entered—and founded—the modern world not as enemies but as associates, or, rather, as identical situations. Such a radical change as overthrowing the "feudal" order required the association of these two powerful values. The principle that "men are free and equal in rights" (the 1789 Declaration) transmutes dominated subjects (and their masters) into equal and free citizens.

The absence of the relation of subjection and domination is, indeed, in a society, both the most basic equality and the most basic liberty. Relations are more intrinsic to society than comparisons are, and, in a relation, freedom from the other's command and equality are practically synonyms. Domination is a person's power to compel another do something, notably by force or threat. By nature, the corresponding subjection is the most vicious of unfreedoms since, in it, a person's will determines another's acts. It is in essence worse than a simple constraint, not only because of the a priori uncertainty, but, much more basically, because it constitutes a kind of amputation of part of the dominated self, and this substitution of wills, this occupation of the other's command center by force (or ruse), is the annihilation of the condition for agency, autonomy, self-respect, and dignity. Domination is usually maintained by force, but it may be worse when the subject endorses the situation in "voluntary serfdom" as Etienne de La Boétie put it. The situation admits of degrees, however, depending on possibilities and costs of avoiding the domination. Slavery is one extreme, and there are many forms of it. Avoiding subjection is sometimes prevented by a status of lower caste one is born in. Serfdom of diverse types also exists, as do lifetime servants of the same master. Domination sometimes masquerades as free exchange, which is fictitious when the alternative is starvation or the lack of satisfaction of some essential need. The wage relationship differs from an exchange of services by its being subjection to the boss's orders within some limits, and the wage earner may have no real alternative or, perhaps, has the only choice to replace one boss by another. This limited possibility to leave the relation also results in a low wage, hence inequality in this respect too, and situations of unequal exchange and exploitation. Intrafamily domination and emancipation toward equal status, power, and rights and duties are major problems of humankind. The domination can also be group-wise, as with colonial situations, and equal status obtained by independence or liberation. All this covers, of course, a large variety of situations according to cases, places, and historical periods.

The absence of subjection, or of strong forms of it, is jointly an equality in itself, relational, and, if all members of a group (or of mankind) have to be free from the corresponding domination, an equality *of* liberty and an equality as generality (or universality).

Equality From Logic

Overview

The basic property of *equal treatment of equals in the relevant characteristics* results from logic for two different reasons. In one, *equality as rationality*, it results from rationality in the relevant and most common sense of providing a reason—justifying. This holds whatever the reason, and even from simply being favorable to provide a reason since the equal treatment of equals turns out to be a necessary condition for all reasons. The second way in which logic requires equal treatment of equals is the property—shortly explained—of "permutable treatment of equals" plus the requirement of full determination (uniqueness) of the result. However, this equality of equals is sometimes an inferior solution and, then, "permutable treatment" is the second-best logically egalitarian concept. The next two sections discuss these topics. Then, the famous principle of "insufficient reason" for inequality is examined and shown to be either fallacious, tautological in two possible ways, or any of the two above reasons.

The relevant characteristics may include, notably, a description of the relations to possibilities. At any rate, this equality can be *prima facie*, that is, in the absence of an overriding reason that may be impossibility or the joint relevance of some other value (which may be the ideal equality of something else, a unanimous benefit from leaving equality, and so on).

Equality From Rationality

Equal is rational, rational is equal. Indeed, rational, in its most common sense, used here, means to give a reason, to justify, or to begin to do it or at least to intend to. It opposes the irrational, unjustified, or arbitrary. Assume individual i receives x_i of the relevant item of any nature (goods, income, wealth, position, right, freedom, power, respect, honor, reputation, consideration, bundles of these, interpersonally comparable level of satisfaction, etc.; the item may even be a rule providing something to an individual as a function of some facts possibly including some characteristics of hers, and the equality is that the same rule is used for various persons, a derived *rule-equality* that will shortly appear to be the very form of rationality itself). If this specific x_i is intrinsically justified, given a reason for, this reason a priori refers to a number of relevant characteristics of individual i, of any nature. The set of these relevant characteristics is denoted as y_i. The reason that leads to choose x_i because of y_i is described by a function

$$x_i = r(y_i). \tag{1}$$

Note that we write equation (1) rather than $x_i = r_i(y_i)$ with a function r_i proper to individual i because, in this case, the reasons, a priori proper to individual i, that leads one to write r_i should be included in the set of relevant characteristics y_i and the function takes form with equation (1). Moreover, a complete social choice determines a unique x_i, and then r is a proper function. Then, if another person, j, has an identical (equal) set of relevant characteristics, $y_j = y_i$, relation (1) implies that she receives $x_j = x_i$. This equality is derived from the simple requirement of justifying, giving a reason, that is, from rationality applied to this social issue.

Note that this rationality provides, in fact, two (equivalent) types of equality: a *conditional equality*, $x_i = x_j$ if $y_i = y_j$, and a *functional equality* meaning that the same function r is used for all individuals, which manifests the universality of rationality (giving a reason) fully applied. The former is also *substitutability*, that is, if another individual j than i, for which $y_j = y_i$, is substituted to individual i, then $x_j = x_i$. The latter is also called *rule-equality*, that is, the same rule r, rather than specific rules r_i possibly different for different i, relates y_i to x_i; rationality (in this most common sense) implies rule-equality. (The converse is not true, although it generally holds; most rules describe reasons; logically, however, there can be rules not justifiable from a "reason.") The equality $x_i = x_j$ results from a requirement of rationality when $y_i = y_j$ (if direct comparisons are furthermore introduced, function r may also depend on x_j for $j \neq i$ for describing these comparisons; then, it should also depend on y_j , and $y_j = y_i$ entails the comparison between x_i and x_j which favors $x_i = x_j$).

In this pure rational equality, there is no direct comparison between x_i and x_j. Sentiments of justice or fairness refer in particular to the choice of the relevant characteristics y_i. This choice implies the answer to the question "equality among whom?" A particular form of characteristics y_i is simply "belonging to a certain set of individuals I"; then the x_i of all these individuals should be *prima facie* equal ("equality from generality").

The property of equal x_i for equal y_i holds *irrespective of the specific reason r*. The simple fact of giving a reason, justifying, suffices for this result. This is common grounds of all reasons and a necessary property for the existence of a reason. Hence, the mere a priori posture or intention to provide a reason whatever it is suffices for the result "$x_i = x_j$ if $y_i = y_j$." This is strictly minimal rationality. A reason that yields unequal results ($x_j \neq x_i$) is applied to different sets of relevant characteristics ($y_i \neq y_j$).

A remarkable consequence is that if one has to share something perfectly divisible between a number of persons who have no other relevant different characteristic, their y_i is—or amounts to—belonging to this group and hence is the same for all, and general a priori rationality (and more generally any particular rule consistent with the constraint) requires equal sharing. No reason can give another choice: any other choice is necessarily without a rule and hence without a reason—that is, irrational. Equal sharing is the only rational (and rule-following) solution (a unique one if all the good is distributed). This is, of course, what is usually done. An example can be drawing lots between these persons: rationality requires allocating equal probabilities to them (actually, this happens to be the basic Condorcet-Laplace axiom of the theory of probabilities).

Permutable Treatment of Equals

Denote as $z_i = (x_i, y_i)$ the pair of x_i and y_i. Choose the set of characteristics y_i as being sufficiently encompassing for z_i to include all that concerns person i for the judgment under consideration. Then, if individual i is attributed z_j instead of z_i whereas individual j is attributed z_i instead of z_j, the two social states are not relevantly discernible and are equivalent for this evaluation. Hence, any permutation of the z_i between the persons i creates equivalent social states. Consider now that all the individuals i belonging to a subset I have the same sets of characteristics $y_i = y$. Then permuting the $z_i = (x_i, y)$ between persons i of I is identical to the same permutation of the x_i only between them. Hence, these permutations of the x_i give equivalent social states. This is *permutable treatment of equals in the relevant characteristics*. Consider three applications of this property.

Full Determination

If some of these x_i differ from one another, these permuted social states are not all identical since at least one individual has different x_i in some of these states. However, a virtue of a principle of social choice is that it be *complete*, providing *full determination*, that is, it designates one of the alternative social states only rather than several equivalent ones. Indeed, notably, action and implementation are the realization of one of these mutually exclusive possible alternatives only, and the principle fully plays its role of guiding the choice solely if it has this property.

Now the states derived by the permutations of the x_i between the persons i of I (with $y_i = y$) are one and the same state if and only if all these x_i are the same. This is equal treatment of equals in the relevant characteristics. Therefore, *permutable treatment of equals plus full determination implies equal treatment of equals.*

Permutability as Second-Best Equality

However, it may be that, actually, some unequal treatment of equals is better than equal treatment of equals. For instance, some collective tasks are better performed with a hierarchical organization of the people, even if they a priori have the same capacities. This is conspicuous for the military defense of society, but it is also the case for many productive or administrative tasks: firms and administrations have everywhere a hierarchical organization. Then, people have different powers, which usually entails different statuses (and unequal pays). Society (and all its members) may also benefit from a differentiated education of people, even if their abilities in all respects are a priori identical. Savings provides another example. Aggregate savings become investment and provide growth. Since people usually save a larger fraction of their income when this income is larger, an unequal distribution of income provides higher aggregate savings even if people have the same propensity to save (as a function of their income), hence a priori a higher growth rate. For a similar reason, private support of the arts benefits from unequal income distributions (with rich sponsors). There may also simply be a limited number of nondivisible consumption goods or tools, and it is usually better that they be actually distributed and used rather than not using them at all, which is the feasible equality. In all such cases, unequal treatment of a priori equals is generally better than possible equal treatment. It may be that everybody benefits from it.

In such situations, permutations of the different x_i—ranks, education, incomes, or items—between individuals *i* with identical relevant characteristics $y_i = y$ provide social states that cannot be judged otherwise than "equally good" from an external standpoint although they are not so for each individual. This permutable treatment of equals is the "egalitarian" property of such cases. The property it keeps from equal treatment of equals is the equal social value—in some sense—of permutations of individuals' allocations. It is a kind of second-best egalitarianism. The drawback is that the corresponding social choice is no longer fully determined by the problem alone, since one of the socially equivalent permuted states has to be chosen. A strictly egalitarian desire to equalize the individual situations leads to an overall worsening. Using lotteries or rotation are classical means to face such situations (both were used, for instance, by the Athenian democracy to fill official positions). Lottery provides a choice with the possibility of ex ante equality, but it leaves the actual, ex post, inequality.

The Principle of Insufficient Reason

Answering the question "Why equality?" by the trivial "Why not?" seems hardly serious. However, "if there is no reason for inequality, choose equality" (or "if there is no good, valid, or sufficient reason for it") is the "reason" for equality proposed by innumerable people, including some of the best minds from Aristotle in the *Nicomachean Ethics*, Thomas Hobbes in *Leviathan*, John Locke in the *Second Treatise on Government,* and Condorcet in his 1789 proposal for a Declaration of Rights, to many contemporary thinkers, including Rawls, who proposes that a good reason would be that everybody or the poorest benefit from inequality, and so on. Consider, however, the following properties of this most famous position for equality.

1. Indeed, if there is no reason, or good or sufficient reason, for inequality, what else can one advocate but equality? This seems to be a tautology about providing reasons. Any other choice would be irrational or arbitrary.

2. However, if this argument in favor of equality is of any use, this implies that there is no other sufficient reason for equality either. Then, consider any state with inequality. There is no reason for any other state, with equality or inequality. Therefore, the same argument leads one to advocate this specific unequal state. Finally, this argument leads one to choose any state, equal or unequal. This apparent tautology is in fact worse: a fallacy.

3. The same reasoning is the "principle of nonsufficient reason," which is the basis of the axiomatic epistemic foundation of the theory of probability, introduced by Laplace and Condorcet: if there is no reason for an event to be more likely than another, attribute equal probabilities to them. However, the principle is, in this context, actually an axiom. This suggests that, in social ethics, this statement could just express a "moral taste," an a priori preference for equality. What it adds to just expressing this is open-mindedness: if there is a reason, a fortiori a good or valid reason, and unavoidably a sufficient reason, for states with inequality, one is ready to abandon this preference. However, equality and inequality are a priori unevenly treated: a reason is required for inequality, not for equality. This is a *prima facie* preference for equality. But not a justified one, so far. Why this unequal treatment of equality and inequality, this asymmetrical status of symmetry and asymmetry?

4. However, preferences do or may also intervene for deciding what counts as a good, valid, acceptable, and in the end sufficient reason for inequality, that is, one that can override the choice of equality. Then, the statement just is: "I choose equality if I do not prefer something else to equality." However, this can mean two things, depending on whether preference alone is considered or the necessity of choosing also is. First, we have pointed out that this choice of equality is to be seen as resulting from a preference. Hence, the statement just says: "I prefer equality if I do not prefer something else to equality." This is a strange preference structure that omits indifference. Second, in fact, a choice has to be made between mutually exclusive alternatives. Then, the statement becomes: "I prefer to choose equality rather than anything else if I do not prefer to choose anything else rather than equality," which now is a tautology.

5. Nevertheless, the absence of reason for inequality may also mean two other things. One is that no imaginable reason for an overall allocation gives inequality. This certainly implies that all possibly relevant characteristics y_i—which could a priori be used for such a reason—are identical for all *i*. Then, for any reason *r*, the definite $x_i = r(y_i)$ are also identical for all *i*. This simply is the general a priori equal treatment of equals.

6. In particular, we may have to choose the allocations x_i to the individuals *i* who belong to a certain set *I*, while we have no (other) reason for this choice. Hence, the relevant characteristic of these individuals is only that they belong to the set *I*. This is y_i for all these *i*. Hence, these y_i are identical. Then any reason based on them gives identical x_i for all *i*. Note that, here, there is no a priori other reason either for equality or for inequality.

7. A different type of reason can justify the principle. With sufficient sets of characteristics y_i, permutations of the individuals' pairs $z_i = (x_i, y_i)$ among the individuals are not discernible. Then, if all these y_i are equal (perhaps just for *i* belonging to the set *I*), this permutation is identical to a permutation of the individual allocations x_i only among the individuals. These permutations are therefore equivalent for any evaluation of the social state. If one is a solution, so are the others. Yet, when the x_i are not all equal, some of these permuted states differ from one another since at least one individual receives different x_i. However, a complete social choice consists of a unique solution. Then, this can only happen if the x_i are all equal. Equality results from permutable treatment of equals and the requirement of full determination of the choice. Sharing the cake between two equal individuals in proportions (1/3, 2/3) or (2/3, 1/3) is equivalent in moral terms although it is not equivalent for each individual. For the proportion (1/2, 1/2) only this multiplicity is avoided. Although there are cases in which equal treatment of equals is less good than unequal permutable allocations, the outcome then is not uniquely determined.

Finally, the insufficient reason for equality is either a fallacy, one of two tautologies, or any of the two basic logical requirements of *prima facie* equality.

Equalities Determining the Overall Distribution

The Five Alternative Equalities of Distributive Justice

Besides the equalities protecting against force in nonsubjection, basic rights, and democracy, the most important role of equalities may concern the overall distribution of the resources of society. Equalities are used in many types of relations. Inspired by Max Weber's remarks about people's ethical judgments, Michael Walzer (1983) argues that this is how it should be with equality in each of a variety of "spheres of justice." One sphere, however, is much more important than others in volume: that in which income distribution is determined (especially since various services can optionally be bought with disposable income—i.e. put in the market sphere). This overall distribution of the resources of society through income belongs to the domain of "macrojustice"—which Rawls calls "social justice," but the common use of this term is sometimes more extensive (including, for instance, the question of handicaps). This contrasts with the multifarious issues of "microjustice" specific as regards goods, people, or circumstances, and with issues of "mesojustice" concerned with specific goods but important ones that concern everybody (e.g., education and health).

For macrojustice, five polar theories of the appropriate distribution are classical and important claims. As for all theories of justice, they are characterized by what they hold should be equal. These equalizands are characterized by two aspects. One is their substance (material, currency, metric) such as income or resources, welfare as happiness, or, in an equality from generality, self-ownership. The second aspect is their structure, as with an ideal equality in individuals' income, resource endowment or welfare, or an equal weight in the highest sum of welfare (utilitarianism) or of incomes. Figure 1 shows this overall structure of the issues. The values of liberties, responsibility, entitlement, happiness, needs, deserts, and merits are implicit, as shortly seen.

These five polar equalities of social justice are very different in nature. The most tangible of these equalities is that of incomes. Welfare classically means, in this context, psychological welfare, for instance satisfaction or happiness.

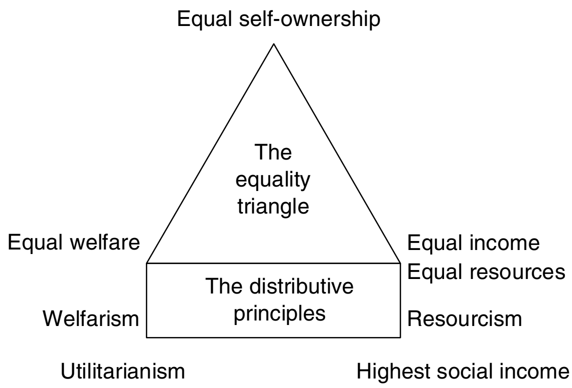


Figure 1 The Topology of Equality

Economists represent it by individuals' utilities. Concepts of equality, addition, or other operations concerning such notions are, of course, problematic, but classical theories consider them and this may more or less provide rough guidelines for policies. Income egalitarians differ from welfare egalitarians by their holding that individuals are accountable for their own different capacities to enjoy (utility functions). If, in addition, people are also entitled to their own capacities to produce and earn (and hence to their resulting income), the result is equal self-ownership. It suffices, for it, to say that each individual has self-ownership—hence it is also an equality from redundancy, generality, or universality.

Equalities in welfare or income from which it is not possible that everyone has more (whereas income transfers are a priori possible) can also be described as maximizing the lowest individual endowment of these items, or "maximin." If equality is desired because individuals who have little of the item have too little, and if some situation with inequality can give more to everybody than situations with equality, equality is to be replaced by maximin (e.g., Rawls's [1971] "difference principle" for "primary goods" and Kolm's [1971] "practical justice" for interpersonally comparable welfare). This assumes that policy can improve the lowest endowment without excessive cost (notably in terms of lowering those of other people).

Equality in weights is a priori anterior—more "upstream"—in the evaluation. Nevertheless, the egalitarian aspect of utilitarianism due to equal weight is classically forcefully (and redundantly) emphasized by Jeremy Bentham and quoted by John Stuart Mill: "each is to count for one and nobody for more than one" (*Utilitarianism,* Chapter 5). It is the basis of Richard Hare's (1981) defense of this philosophy as an interpretation of Kant's view that each individual should be given equal consideration (it seems, however, that the product of individual utilities would not give them less equal consideration than their sum—it amounts to comparing relative variations in utilities rather than their absolute variations).

The Multiple Equalities of the Overall Distribution

If everybody, which includes voters and officials, holds that some social principle is irrelevant for a problem, this principle cannot be implemented for this question on social grounds. Now people actually hold that the comparison of individuals' capacities to enjoy (hedonistic capacities) or their variations, and of their tastes, is relevant for allocative choices in two types of cases: when they refer to suffering and when the distribution is between people who sufficiently know one another to feel empathy towards the others. Allocations in a hospital or in a family are typical cases. If national fraternity actually ruled the minds, or in case of national disasters creating general suffering, the principle of overall national distribution (for instance for the income tax) would be in the welfarist family. In the other cases, people's opinions about income distribution are instances and associations of the other two cases only, income egalitarianism on the one hand, and the self-ownership of classical liberalism on the other. The resulting social and political synthesis or compromise is a mix of these two values.

The normal way of representing the resulting incomes is that they are the sum of two parts, an egalitarian income and a classical liberal one. For clarity, denote as i one of the n individuals, y_i her income, ℓ_i her labor, w_i her wage rate. Her earned income is $w_i\ell_i$. The average wage is $\bar{w}$. The *egalitarian income* is the *equal* sharing of individuals' earnings during an *equal* labor k, kw_i for individual i. This egalitarian income is $k\bar{w}$. Above that, however, individuals are *free* to work ℓ_i and keep their earning from the extra labor, $(\ell_i - k)w_i$. Their disposable income is

$$y_i = k\bar{w} + (\ell_i - k)w_i.$$

For individuals participating in this redistribution, the equalization labor should be such that $k \leq \ell_i$ because people do not accept taxing leisure $k - \ell_i$ at the value of labor (if $w_i > \bar{w}$), and providing a wage supplement (of $(\bar{w} - w_i)$ if $w_i < \bar{w}$) to hours $k - \ell_i$ which provide no wage seems absurd.

This redistribution is egalitarian in various respects. On tangible grounds, it is the more egalitarian the higher the equalization labor k is (it is not at all for $k = 0$, the pure self-ownership of classical liberalism). On rational grounds, it has a number of remarkable egalitarian structures. It transfers *from* each *equally* in labor (or in equal proportion of her capacities), and *to* each *equally* in income. It implies an *equal minimum income* $k\bar{w}$. It amounts to each receiving an *equal universal basic income* $k\bar{w}$ financed by an *equal labor* k of each or in *equal proportion* k of each capacity w_i. It also amounts to *each individual equally yielding to each other the product* $w_i \cdot (k/n)$ of an *equal labor k/n* in a kind of *general equal labor reciprocity*. It is also *equal free exchange* (of labor) *from an equal allocation* ($k\bar{w}$ in income and k in labor or the complementary leisure). The two parts of income are *equality according to deserts* (equal income $k\bar{w}$ for the equal labor k) and to *merit* (i.e. including the effects of personal capacities w_i for labor $\ell_i - k$), respectively. Individual i's leisure is $\lambda_i = 1 - \ell_i$ if total time is measured as 1. Individual i's *total income*, including the value of her leisure, $w_i\lambda_i$, is $Y_i = y_i + w_i\lambda_i = w_i + k(\bar{w} - w_i) = k\bar{w} + (1 - k)w_i$. But $P_i = k\bar{w} + (1 - k)w_i$ is the value of a price index for the two prices of income or consumption goods, 1, and leisure, w_i for individual i, with respective weights $k\bar{w}$ and $1 - k$. Hence, Y_i/P_i is the same for all i. But Y_i/P_i is by definition the corresponding "real income" or "purchasing power" of individual i. Therefore, the distribution in question also amounts to *equal purchasing power*, which is a kind of *equal real liberty*, for individuals' choices of income (or consumption) and leisure (or labor). This equal liberty is not an identity of domains of choice (with different w_i).

Note that basing a tax on the wage rate can be done as in the present French tax law, by exempting overtime labor earnings from the income tax, over a low benchmark. There is de facto no cheating (because it would be too complicated to hide this basis from the possible controls). The full theory adds other dimensions of labor than duration, notably formation.

Equality of Liberty

Basic Rights or Equal Negative, Protective, or Civic Liberty

The use or threat of force may be steady or occasional. A person may incur it from others as individuals, in groups, or through institutions. The absence of such forceful interference defines a freedom called social, protective, negative (a term used by Kant, Mill, and Isaiah Berlin), or civic (Mill). Its application to various specific issues constitutes the basic rights or basic liberties. With this freedom, a forceful constraint on someone can only implement a previous acceptation of it and notably a previous agreement (possibly an implicit one) of this person. This absence of force in inter-individual relations is an equality, and a general basic demand of modern society is that all individuals benefit equally from such liberty (equality as generality or universality). This demand is even that this liberty has priority (this is the first and main statement of the corresponding constitutions).

Is this general equal liberty with priority possible, however? This raises an essential conceptual issue with important consequences. The text of the French *Declaration of the Rights of Man*, and many thinkers such as Rousseau, Condorcet, Mill, and Rawls have held that these basic liberties or rights should be, with priority, "equal for all and, then, maximal" (Rawls even admits inequality if it permits each of the least endowed to have more, as he does about "primary goods"). However, they consider jointly these rights and liberties plus some means to make them actual possibilities or "real" (as Marx puts it). They want to face Anatole France's objection that: "rich and poor people alike have an equal right to sleep under the bridges" (the traditional shelter of Paris tramps). However, there is no a priori limit to these means (to the number of private planes and airports for freedom to move, the size of private cathedrals for freedom of worship, the privately owned media for freedom of expression). Then, this principle uses all the resources of society without even a principle for choosing between these various real liberties. The solution of defining some amount of means for each right is a priori arbitrary. Moreover, any equality of these means for all would be found worse by everybody than some other, unequal solution because people make different uses of these rights and have different preferences about them.

The rational solution consists in distinguishing the formal rights from the means of benefiting from them, and in putting the question of the means in that of the general distribution and of the free exchanges of goods resulting from it (with the possibility of some minimum income). Then, when actions or intentions of different individuals oppose one another and cannot be implemented jointly, this opposition can be attributed to the means and is solved by the property rights concerning them (for instance, the occupation of the same place at some time) and not to the "formal" rights in themselves. Then these rights are nonrival between themselves and can be equally held in full and used at satiety by everybody.

Equal Real Liberty

The next issue about liberty concerns people's means of free action, which was the topic of the section on the overall distribution. The necessary distinction between general "macrojustice" and more specific issues of "microjustice" and "mesojustice" is explained there. The basic liberties imply equal freedom of exchange, given the overall income allocation. The theory of macrojustice obtains a structure of distribution that can be defined in various ways as equal liberty (although with different domains of choice as a result of the different earning and productive capacities of the individuals—see Kolm, 2004, 2008a, 2010).

Equality of Opportunity

Equality of opportunity describes a set of cases of equality of liberty to be found in various issues of justice at all levels. A priori it means the identity, for various agents, of a set of alternatives among which each can choose. In the standard and most common meaning, this refers to social conditions of access to certain benefits, positions, situations, jobs (for instance, with regard to various types of discrimination or family influence), or possibilities (such as access to various types of education). The concept has been extended to transform these formal freedoms into more "real" ones, and these direct choices into their outcomes, by adding the effects of personal capacities and social settings and thus considering opportunities for income, achieved level of education, or the actual performance permitted by jobs or positions (perhaps for given levels of effort). One could also consider equality of opportunity for welfare as happiness or satisfaction by adding the consideration of capacities to enjoy or be satisfied. Equality of opportunity thus describes cases that are different and sometimes opposed. This explains why politicians of all kinds love the concept whereas practically all philosophers criticize it severely.

The initial motives for equality of opportunity arose or arise from two different and opposed perspectives, one for realizing an equality and the other for criticizing another equality. The equality to be realized is that of some possibility of choice, as the name indicates (for example, one wants some people to have access to certain positions without discrimination or other obstacle). The other motive is the objection to policies of equalization or uniformization, for different people, of outcomes due in part to their actions, and the demand to replace this equality by that of domains of choice in which these agents choose, thus replacing an equal outcome by an equal liberty to pursue this end. Since this change generally leads to unequal results of actions, this stance is anti-egalitarian in this sense. The emphasis is often not only on the comparison between the agents but also on some competition between them, for which the equality of opportunity is supposed to provide fair conditions. This elicits the classical leftist judgment, preferring not only the "actual" outcome equality but also or mainly the possibly convivial relationships jeopardized by the competition, as expressed in the statement that "equality of opportunity is good for horse races but not for humans." However, equality of opportunity includes both liberty and equality in possibilities; the absence of such equality includes cases that are generally considered to constitute the most viciously unjust features of societies.

The basic feature and difference between the cases are in the definition of equal versus unequal opportunities. The main one refers to discrimination that limits choices on the basis of such criteria as race/ethnicity, family, caste, feudal-like order, gender, and religion. Apart from formal discrimination, the principle often refers to advantages provided by family relations, including favoritism, nepotism, social networks, information, direct support, and the role of families in education at home or school. A basic issue is whether personal capacities, innate or due to family influence, notably in childhood, are counted among the sources of the opportunities in question. The famous meritocratic slogan "the career opened to talents" refers to a situation in which positions are allocated according to talents alone, banning other social discrimination. Policies described as "positive discrimination" often are attempts to compensate for disadvantages resulting from social setting and family influence.

The simple slogan "equality of opportunity" thus covers a number of cases that are quite different and can belong to opposing ethical positions. Formally, there are several types of equality of opportunity.

Negative equal opportunity in action bans formal social discrimination of all kinds.

Positive equal opportunity in action helps people who cannot perform some relevant action to actually perform it.

Equal opportunity in action and result implies that if some people choose to perform the same action, they will obtain the same outcome, possibly with help for those who are disadvantaged as a result of unfavorable personal capacities or circumstances (notably social environment).

The next step would simply be equality in outcome, which is not equality of opportunity from the point of view of its causes, but which can be equality of opportunity for the further use of the outcome in so far as it is an intermediate product—such as wealth or education or health as used for further choices and actions.

Equality and Responsibility: Responsibility-Free Equality and Equal Joint Responsibility

Responsibility is the assignment of the effects of an action to the actor. It requires freedom, which conversely implies responsibility if this issue is raised. The allocative principle of responsibility holds that this should determine the allocation of benefits and costs (and not only blame or praise). This principle can be equally applied to several or all actors. If this is the only principle, facts of society that do not result from members' actions have no particular reason for their allocation and therefore their value should be equally shared (from rationality). However, beneficial or detrimental aspects of society are generally joint products of the actions of several people and of given facts. Apart from particular structures of the effects of the acts (separability, additivity, symmetry), there is no a priori solution to the corresponding allocation of costs or benefits. However, when several agents' acts and the allocation of the resulting benefits or liabilities are collectively chosen by a required unanimous agreement between these people, each of them has a veto power on the realization; therefore she is fully responsible for it, and hence in particular for her own act and her benefit or liability. When there is some impediment to this general agreement (in the nature of a "transaction cost"), a solution is to estimate what this agreement would have been were this impediment absent and to impose the obtained sharings and acts (this is a "liberal social contract"—see Kolm, 1985). Recently, the responsibility-free equality has been emphasized by Gerald Cohen (1989), and closely analyzed on philosophical grounds by Matt Matravers (2007) and on economic grounds by Marc Fleurbaey (2008). A particular important application on a delicate point differentiates two main theories of what should be equal. Rawls (1971) and Ronald Dworkin (1981) hold people fully responsible for their own tastes, preferences, capacities to enjoy (and hence "utility functions"), which is a priori rather strange but can be defended by considering people's ends, desires or values as "ambitions" or "life plans" chosen by them. In contrast, the ordinary "welfarist" theories include compensations derived from differences between individuals' hedonic capacities.

Fundamental Insurance: Equal Hypothetical Liberty

A "fundamental insurance" is a hypothetical mutual insurance undertaken by people against the risk of some disadvantage that in fact they already have. This may be, for instance, having poor health or having received a poor education because of family status. This theory provides a rationale for corresponding compensating transfers that mitigate the inequality. It rests on a putative free choice (an exchange, which makes it be a kind of "social contract" in the technical sense), and it also is a "partial original position" with a "partial veil of ignorance." Its assimilation of a choice concerning justice to a choice in uncertainty is a priori problematic but is to be accepted if this is general opinion. This is the case, for instance, for the European system of public health insurance; the fact that what people pay does not depend on their given propensities to be sick implies a "fundamental insurance" of these handicaps. This particular kind of "liberal social contracts" (the impediment to the insurance agreement is the "arrow of time") has also been directly suggested by Dworkin (1981).

Comparative Egalitarianism

Sentiments favoring equality are often the result of direct, intuition-like comparative judgments. However, the logical "equal treatment of equals in the relevant characteristics" certainly lurks behind such emotions. Nevertheless, such judgments seem close to aesthetic ones, as remarked by Kant and confirmed by location on brain imagery (*fair* comes from a Germanic word referring to beauty, and the Greek and Latin concepts of beauty, *kalon* and *pulchrum*, were never neatly and consistently distinguished from the moral good).

When the individual items that are compared have several dimensions about which the individuals can have different preferences (this is the case, for instance, for bundles of commodities), it is possible that *no individual prefers any other's "allocation" to her own* without these individual allocations being identical. This principle, called *equity-no-envy*, is one of the most commonly used in egalitarian studies since the early 1970s. Its egalitarian properties are readily seen. If there is one (desired) dimension only, the principle implies equality. If the individuals have identical preferences, the principle implies that they are indifferent between all individual allocations. However, the most important egalitarian property of this principle is that it amounts to an equality in liberty. It is satisfied if and only if there exists a domain of choice such that each individual's allocation can be chosen by this individual (with her given preferences) in a domain identical to it. The analysis of this principle in Kolm (1971), after mentions by Jan Tinbergen and Duncan Foley, was followed by a large number of applications and variants reviewed by William Thomson (2008). It cannot be called "no envy" by itself because the sentiment of envy arises from the joint presence of the other's and one's own allocation in one's "utility function," but it is formally related to structural properties of a genuine theory of envy (Kolm, 1995).

Political Equality: Democracy

One of the most important applications of equality is to politics, in the realm of democracy. The Athenian four equalities of democracy still provide the basic framework:

> equality in the eyes of the law, or *isonomia;*
>
> equality in voting, "one man one vote," or *democracy stricto sensu;*
>
> equality in public expression for influencing others, or *isegoria*, applied in Athens as a right to equal time of speech in the assembly of citizen;
>
> equality in the *access to official positions*, implemented there by drawing lots or by rotation.

This was for a middle-sized society, with officials but a priori the possibility of mutual influence between citizens. Women, slaves, and foreigners were excluded, and official positions soon became the privilege of members of influent families.

Later democracies had a variety of restrictions to voting rights or access to positions. In present-day mass societies, the most violated democratic equality is *isegoria*, since the flow of public messages is that of the mass media in which a tiny aristocracy of journalists, media owners, and politicians send views, values, information (and entertainment) to the mass of the people who are gagged in this respect (the Internet opens new possibilities that already has had some important political effects, but it is not clear whether this will or can develop at the level of the overall problem). Although democracy has led to important inequality-lowering redistributive transfers in many countries, it also allows corporate and other interest groups to buy favorable laws with contributions to campaigns. The various specific rules of particular democratic systems also raise innumerable issues with respect to equality. In the end, the main egalitarian virtue of democracy is as a barrage against dictatorship, the harshest inequality.

Equality and Impartiality

From a social point of view, humans manage to shelter two opposed selves "in their breast" as Adam Smith puts it. Their self-centered and partial self favors themselves and the people they like or have particular relations with only. However, they also have an impartial self able to take an objective view putting everybody—themselves included—on the same footing: This is what Thomas Nagel (1986, 1991) calls "the view from nowhere"—but is there such a place as nowhere?

A priori, the distribution most favorable to someone's strict interest is without equality. Moreover, among situations with equality of some kind, one (in general) is most favorable to this person's interest than the others. For instance it is equal incomes (barring incentive effects) for people with low earning power, and equal self-ownership in the free market for people with high earning capacities. The person's impartial self, however, will generally make some other choice with some sort of equality. But will this latter choice be the same for everybody? This is often believed, for instance, by Adam Smith; by Mill, who takes the equality to be that of the weights in an egalitarian sum of utilities; and probably by Thomas Nagel. However, the a priori only logical structural requirement of an impartial judgment is that it respects equality of some sort. Hence, there is a priori a very large choice for such a judgment. Moreover, the impartial individual evaluation uses some psychological characteristics of this person that a priori differ from one person to the other. Therefore, there tends a priori to be different impartial evaluations (hence with different equalities) for the various individuals—other phenomena such as a common moral culture or mutual influence through dialog can change this.

However, since an individual's impartial judgment of society does not depend on her own specific interests, she produces such a judgment if she is asked to judge when she does not know which specific person she is, which specific situation is hers. Each individual may still be purely self-interested, but her ignorance prevents her from favoring her own actual self-interest. On can think of this as a time sequence, with evaluations "before" the individuals know which specific person they will be, before the actual individual situations—with the specific interests attached to them—are assigned. Rawls (1971) calls this state the "original position," in which individuals evaluate "behind the veil of ignorance," and this way of thinking is the most famous modern theory of impartiality. A first result appears straightforwardly: Since the individuals in the original position face the same (uncertain) prospect, their evaluation there are the same: They make an equal, *unanimous* "original social choice." Rawls then assumes that some possible risks faced by these individuals are so severe that they choose maximal protection thanks to three principles in order of priority: (1) the basic liberties (equal and maximal); (2) nondiscrimination; (3) the "difference principle," that is, the highest possible level of the lowest individual endowment of "primary goods" (one of which is income or wealth), which is not an equality because incentive and disincentive effects permit unequal allocations to induce a higher production, hence to provide more to the least endowed. This implies, however, that the individuals in the original position, who are assumed to make a social contract for the adoption of the "principles of justice," do not also agree not to follow their self-interested impulses when working in real life.

Rawls built this theory in order to provide an alternative to utilitarianism, then the prevailing view in English-language philosophy. However, John Harsanyi (1976) produced another theory of the original position, which leads to a utilitarian-like structure! He finds that the individuals in the original position want to maximize a sum of individuals' utilities obtained from the mathematical expectation of utility with equal probabilities of being all the actual individuals. Yet the various individuals, in the original position, should in fact have different maximands because they a priori differ with respect to (1) their preferences for being the actual individuals (some prefer to be rich, others prefer to be liked or famous, still others prefer to be good), and (2) their risk-aversion. For both theories, moreover, the reduction of a choice of justice or social optimality to a self-interested choice in uncertainty is problematic: the former is accountable to society and ethics, as the second is not. The individuals may also be risk-seekers, which leads to the most unegalitarian result. The theory of "moral time-sharing," which asks the individuals to consider that they are the various actual individuals successively in time is another impartiality theory, which, however, raises similar problems.

Equality as Lower Inequality

Philosophers discuss equality, but since large equalities do not exist in real life, sociologists study inequality (Melvin Tumin, 1967, provides a good overview) and economists compare and measure inequalities. This comparison and measure of inequality has developed into a very large field of studies since the mid-1960s. Various questions are analyzed: Does income inequality increase or decrease when all incomes vary in the same proportion or by the same amount? Does a transfer from a richer person to a poorer one diminish inequality (it augments the inequalities between the poorer and the equally poor and still poorer, and between the richer and the equally rich and still richer, yet it may be favored on the grounds of welfare if the poorer suffers from this poverty)? Are the relevant inequalities relative or absolute? And so on (Kolm, 1966). Multidimensional inequalities and inequalities in liberty are also studied. The *Handbook of Income Inequality Measurement*, edited by Jacques Silber (2000), gathers reflections of most of the experts.

Multidimensional equalities, that is, equalities in each of several goods, are in general such that other, unequal distributions are preferred by everybody, because people a priori have different tastes. However, among allocations of these goods that are not so dominated, some can be defined as "more equal" than the others. But if each individual consumes some of each good, these solutions amount to equal incomes (Kolm, 1977, 1996b). One famous proposal of such multidimensional equality is Sen's (1985) for individuals' "capabilities"; the noted result applies to it.

Positive Relational Equalities, Reciprocity

Equality in social relations does not solely involve the noted absences of domination, too unequal distribution, or envy. It has many other dimensions. Equality can also be with respect to status, respect, and consideration, with, notably, mutual respect and consideration. In such a society, people relate to one another on an equal footing and interact with others as *alter ego*. They are knights of the round table of society. Such a *society of equals* is something other than an egalitarian society, although it certainly limits inequalities of various types. It adds a requirement of liberty in the consideration of others, which situates these relations on the verge of fraternity. Relations between equals are described in particular by Marcel Mauss (1924) and David Miller and Wayne Norman, in Mason (1998).

These positive relational equalities can be supported by a basic sociopsychological property of humans: the tendency to treat others and relate to them as they treat you and relate to you; such reciprocity is a relational egalitarian reaction that is a main cement of society (Kolm 1984, 2008b).

Serge-Christophe Kolm
École des Hautes Études en Sciences Sociales
Paris, France

See also Inequality, Economic; Inequality, Political; Justice

Further Readings

Barry, B. (1975). *Justice as impartiality*. Oxford, UK: Clarendon Press.

Berlin, I. (1955–1956). Equality. *Proceedings of the Aristotelian Society, 56,* 301–326.

Cohen, G. A. (1989). On the currency of egalitarian justice. *Ethics, 99,* 906–944.

Cohen, G. A. (2008). *Rescuing justice and equality*. Cambridge, MA: Harvard University Press.

Dworkin, R. (1981). What is equality? Part I: Equality of welfare. Part II: Equality of resources. *Philosophy and Public Affairs, 10*(3/4), 185–246, 283–345.

Dworkin, R. (2000). *Sovereign virtue: The theory and practice of equality*. Cambridge MA: Harvard University Press.

Fleurbaey, M. (2008). *Fairness, responsibility and welfare*. Oxford, UK: Oxford University Press.

Hare, R. M. (1981). *Moral thinking*. Oxford, UK: Oxford University Press.

Harsanyi, J. C. (1976). *Essays in ethics, social behaviour and scientifc explanation*. Dordrecht, Netherlands: Reidel.

Kolm, S.-Ch. (1971). *Justice et équité*. Paris: CEPREMAP. (English translation, *Justice and equity*, 1998, Cambridge: MIT Press)

Kolm, S.-Ch. (1985). *Le Contrat Social Libéral* [The liberal social contract]. Paris: Presses Universitaires de France.

Kolm, S.-Ch. (1996). *Modern theories of justice*. Cambridge: MIT Press.

Kolm, S.-Ch. (2004). *Macrojustice: The political economy of fairness*. New York: Cambridge University Press.

Kolm, S.-Ch. (2008). *Reciprocity: An economics of social relations*. Cambridge, UK: Cambridge University Press.

Mason, A. (1998). *Ideals of equality*. Oxford, UK: Blackwell.

Matravers, M. (2007). *Responsibility and justice*. Malden, MA: Polity Press.

Mauss, M. (1924). Essai sur le don. *Année sociologique n.s., 1,* 30–186. (English translation, *The gift*, 1967, New York: W. W. Norton)

Miller, D. (1999). *Principles of social justice*. Cambridge, MA: Harvard University Press.

Nagel, T. (1991). *Equality and partiality*. Oxford, UK: Oxford University Press.

Parfit, D. (1995). *Equality or priority* (The Lindley Lectures). Lawrence: University of Kansas.

Pojman, L., & Westmoreland, R. (Eds.). (1997). *Equality: Selected readings*. Oxford, UK: Oxford University Press.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Rousseau, J.-J. (1984). *A discourse on inequality* (M. W. Cranston, Trans.). London: Penguin. (Original work published 1755)

Silber, J. (2000). *Handbook of income inequality measurement*. Dordrecht, Netherlands: Kluwer.

Thomson, W. (2008). Fair allocation rules. In K. Arrow, A. Sen, & K. Suzumura (Eds.), *Handbook of social choice and welfare* (pp. 393–506). Amsterdam: North-Holland.

Walzer, M. (1983). *Spheres of justice. A defence of pluralism and equality*. New York: Basic Books.

Equality, Political

Political equality is understood as a condition in which all citizens have equal influence on the collective decision making of a political community. Political equality can never be fully realized. It stands in tension with social and economic power as well as cultural hegemony, but countries and communities differ greatly in the degrees and forms of political inequality. Democracy puts a premium on reducing political inequality, and for many the measure of political equality points to a critical dimension along which the quality of democracy varies. Yet there are significant controversies on whether and how public policy should try to reduce political inequality. Whatever one's normative position, it is important to recognize the reality of major disparities in the power resources of "the few" and "the many." These disparities explain the difficulties of coming closer to political equality. Sheer numbers aside, collective organization is the main form of power available to the many, counterbalancing the advantages of dominant groups.

During most of recorded human history, claiming political equality as an ideal was an incomprehensible notion, except for a few who retained in subordinate positions sufficient independence of thought. Political equality became conceivable as a goal with the rise of democracy. Instigated by excluded groups who increasingly could make their voice heard, even rather elementary forms of democracy created a zone of relative autonomy within which political decisions are taken. Democratic politics must to some extent be insulated from the overall structures of power and the system of social inequality as a whole if citizens are to be able to exercise democratic freedoms. Conventionally defined by freedoms of expression and association, regular elections with comprehensive suffrage, and the government's responsibility to those elected as the people's representatives, a democratic polity gives citizens equal political rights—to vote, to equal protection of political freedoms, and to equality before the law. Critics have deprecated the conventional definition of democracy as merely "formal" because it works to even out advantages only within a circumscribed political sphere while overlooking broader and subtler ways in which social inequality shapes politics. Yet even formal democracy opens the possibility of political contest about the acceptable degree of political inequality.

Dominant groups can use their social and economic power resources in the political sphere. Specifically, they can use their influence over education, cultural productions, and mass communications—their cultural hegemony—to shape in less than obvious ways the views, values, and preferences of subordinate groups. If these effects of social and economic equality are not substantially contained, political equality will be extremely limited.

The distance between the ideal and the reality of democratic equality has varied greatly over time and continues to vary widely from one country to another. In each instance, this equality gap represents a compromise between dominant groups and the many, often shaped by political and social institutions of long standing. The underlying factors include the power balance within society, the relations between the state and civil society, international power constellations, the organization and the degree of cultural autonomy that subordinate groups enjoy, and the extent to which dominant groups see democratization as a threat to their interests.

Competing Concerns

Advancing political equality is clearly not an uncontested goal. It is at odds with major interests, and it may involve the sacrifice of competing values. Opposition to greater political equality, whether based on interest or principle, is likely to vary depending on whether the topic is the political realm narrowly conceived, the conversion of socioeconomic conditions to political advantage, or the overall impact of social inequality on the political sphere.

Does increasing political equality constitute an unqualified good? Or does it involve sacrificing other important values? If there are trade-offs, how do we judge different outcomes? These are normative issues, but their discussion often intertwines the empirical with normative claims. Thus, it seems reasonable to sketch some of these contested issues and to indicate alternative views of them.

Even if we stay purely within the political realm, there are two major objections to promoting political equality as much as possible. The first is that "the many" are not competent to choose reasonable policies. The second is that increasing the number of participants can worsen coordination problems to the point of system overload and ungovernability. The conventional form of democratic rule—representative democracy—answers most of these objections by interposing a professional state apparatus (responsible to elected officials) between voters and policy outcomes, though the issues of voter competence and coordination problems will vary widely across countries and historical situations.

Both the objection from competence and the objection from numbers raise issues of reduced efficiency and its costs. These seem outweighed in principle by three considerations of a different sort. First, "efficiency" takes on substantive meaning only in light of the aims being pursued, and these in turn depend on material and immaterial interests on which people are often divided. Second, a reasonable principle holds that all members of a political community whose interests are affected by collective decisions should have a say in them; affected interests trump limited competence. Finally, political equality is an acknowledgment of the decisively similar dignity of all citizens as human beings who are entitled to the rule of reciprocity. In practice, broader measures aimed at reducing political inequality—among them high-quality general education, multiple sources of information and analysis, and trustworthy organizations to represent interests that are otherwise at a disadvantage—may in effect also cut these costs in efficiency.

Limiting the conversion of assets from other spheres into political advantages is far more controversial. And objections mount even further against policy proposals to level differential assets outside the sphere of politics because spillover effects are too difficult to contain.

There is first the issue of using economic assets in electoral campaigns. In *Citizens United v. Federal Election Commission* (2010), the U.S. Supreme Court has ruled, much to the outrage of some, that spending money in campaigns for political office enjoys the same protection as free speech. Critics of the decision note that the Court, in effect, dealt with economic inequality as if it were irrelevant to the integrity of the democratic political process. While there may be other ways of evening out the financial aspects of electoral competition, none is likely to shut off the flow of money into politics completely.

Economic inequality has consequences for political equality that go far beyond the issues of a level playing field in electoral campaigns. The wealthy can disproportionately influence how policies are made and implemented. Monopolistic and oligopolistic market power can easily be turned into political bargaining advantage. Different levels of government—from local to national—are dependent on investment decisions of large corporations. In addition, most voters in industrial and knowledge-based societies depend for their economic security on employment. The threat of unemployment is taken for granted in capitalist societies, but it has a major impact on their political dynamics. All advanced capitalist countries have legislation that seeks to constrain the creation and use of monopolistic economic power. But such measures have at best slowed the trend toward concentration of corporate ownership, and few have focused on blunting their impact on politics. Measures to soften the impact of unemployment enjoy widely varying support across countries.

Attempts to go further in limiting the conversion of economic power into political gain are subject to fundamental ideological objections as well as to pragmatic arguments that the pursuit of equality will harm economic growth. While the value of economic freedom for entrepreneurs and capital owners is in principle hardly a match for the value of increased political equality, claims that measures compromising the functioning of markets also impair growth have persuasive power in Western countries. The more comprehensive welfare states in advanced capitalist countries insist, however, that this protection of the market mechanism has to be combined with relatively generous compensatory policies, which make the functioning of the market socially sustainable.

If measures to limit the conversion of economic advantage into political power have only a partial effect, does the pursuit of political equality warrant policies that reduce inequalities of income and wealth more directly? There can be little doubt that differences in the distribution of income and wealth across countries and over time within countries make for significant variations in political equality.

Competing value claims dominate the discussion of policies meant to change the distribution of income and wealth. It is on the issue of direct political action to reduce income and wealth differentials that the claims of an inherent contradiction between equality and freedom have concentrated. These claims are plausible when one thinks of sudden policies of increased taxation and expropriation that are aimed at leveling incomes and wealth. Such policies would be fought by the privileged with all means at their disposal, potentially endangering democratic rule. In turn, breaking this opposition could indeed lead to ruthless dictatorship. Reduced levels of economic inequality would then coincide with the destruction of democratic equality. On the other hand, we know of examples—the Scandinavian countries are the most prominent—where popular policies pursued over decades have resulted in significant reductions of after-tax and after-transfer income inequalities. Yet political freedom remained unimpaired, and no long-run drop in economic growth has accompanied these policies.

The other major issues beset by value controversies concern cultural hegemony. There is little question that relatively small elites have a disproportionate influence on the production of culture as well as on its diffusion through education and the mass media. Yet this influence is in most advanced capitalist countries quite heterogeneous in character. It is pluralistic; but it is unequal nevertheless. Cultural hegemony, however manifold and complex, creates substantial political inequality.

Policy responses to issues of cultural hegemony are limited. Apart from some attempts at hindering if not obstructing the trend toward concentrated ownership of the media and support for public education and public broadcasting networks, the standards of news journalism and of academic freedom and analytic universalism constitute an important protection against direct conversions of wealth and political influence into control over news and research.

At the same time, measures that would go beyond these broadly accepted policies of insulating certain cultural institutions and spheres against direct partisan influence face nearly insuperable difficulties. Most people who care about undue influence in cultural production and diffusion are part of the dominant, though internally heterogeneous, mainstream. Efforts to curtail such influence, furthermore, quickly run up against two important and commonly accepted principles—freedom of expression and respect for the autonomy of different cultural spheres such as art, news reporting, or academic research. Furthermore, since some protections against the conversion of economic and political clout into cultural influence rely precisely on these principles, policies at odds with them would at least partly be counterproductive.

Power Resources and Political Equality

If competing values argue against reining in processes that have inegalitarian implications, the effects of these processes are not thereby rendered any less real. A clearer picture of the chances of decreasing, maintaining, or increasing political inequality can therefore benefit from a brief overview of the main power resources of the few and the many.

Coercive power is a fundamental resource that is inherently distributed unequally. If the state succeeds in monopolizing coercive power—as most states strive to do—democratic equality is protected only if the use of that power is regulated by law and if equality before the law is sufficiently realized to rule out differential intimidation for political gain. While this can roughly be taken for granted in most rich democracies today, pockets of private coercive power and gross imperfections in equality before the law are major factors subverting political equality in a number of countries in the developing and postcommunist worlds.

The state *administrative apparatus* represents another power resource of great reach, combining expertise, tax-based funding, readymade organizational capacity, and ultimate recourse to coercion. The administrative organizations of modern states preempt a good deal of decision making and can create substantial imbalances in political equality. Supranational administrative bodies such as the European Union or the World Trade Organization wield great power unchecked by democratic controls.

Economic advantage is a major obstacle standing in the way of realistic political equality. Capital ownership, as already indicated, carries great power, the more so the greater the mobility of capital. Again, transnational corporations, much as transnational public agencies, wield power that is unconstrained by democratic controls, except in conditions of severe crisis.

Income differences are less concentrated than capital ownership, but they vary greatly from one country to another as well as across time periods. The two ends of a distribution are most significant for democratic equality. The poor never attain influence commensurate with their numbers, and the very rich nearly inevitably have disproportionate political clout.

Social status represents more than a point on a scale of esteem; it involves social attachments and aversions. It shapes interaction patterns, offering entrance to, and imposing exclusion from, different social circles. Politically, it defines the chance to be heard and to be trusted; it increases or diminishes one's political "voice." The importance of status differentials varies considerably across countries, but status hierarchies tend to persist over time. However, some major correlates of status, including economic position and level of education, are not policy independent. Broadening access to advanced education and flattening the distribution of posttax and posttransfer incomes reduce even persistent differences grounded in ethnic, racial, and gender status.

The unequal influence on the production and diffusion of culture exerted by varied but small cultural, religious, and economic elites constitutes, as noted, a major problem for political equality. Cultural hegemony represents inequality of a peculiar kind: Its effects remain largely invisible to those whose views, values, and preferences it shapes, and they are taken for granted by those whose material and immaterial interests find expression under the hegemonic status quo.

Many regard arguments about cultural hegemony as spurious expressions of Marxist ideology—the claims of "false consciousness" in a new guise. But the conception of cultural hegemony sketched here does not claim superior knowledge of people's objective interests; it simply argues that elite views and interests tend to shape the views and preferences of subordinate groups more than vice versa.

It is useful to explore these problems of cultural inequality in regard to the specific issue of politically relevant knowledge. Political knowledge is a major power resource, and its distribution is profoundly unequal. This is true most obviously in regard to information. Those who are favored by education and their position in networks of information have clear advantages over those less favored. Furthermore, for judging and absorbing the flow of information and for the ability to resist spin, complex background knowledge is necessary. This is largely a function of education and of the repeated exercise of such assessments.

Beyond these cognitive advantages, there is another more basic kind. It concerns the knowledge that is generated by systematic inquiry. Here, the question is not how this newly generated knowledge is disseminated but whether the problem formulations and specific questions asked in research are shaped by concerns and presuppositions as well as by blind spots that correspond to the concerns and blind spots of select groups while neglecting the interests and perspectives of others.

The establishment of feminist and African American scholarly niches may have gone against the norms, but it certainly is an indication that for a long time these norms of universalism had rather disappointing results. The problem may well be equally serious for class-based cognitive interests. Researchers from a lower class background acquire a new social and economic status in the process of their training and career, and they have better career chances if they stay within established theoretical frameworks of basic assumptions and problem formulation.

That the orientation of research is of great importance for long-term political gains is also reflected in the proliferation of privately financed (though tax supported) think tanks in the United States, many of which have a right-of-center outlook. In many other Western countries, the orientation of similar centers is less weighted toward one side and often balanced by publicly supported institutions.

Organization for collective action in voluntary associations, unions, and parties is the most promising power resource of "the many." Collective organization can compensate for the impact of social and economic inequality on democratic politics. It can mobilize voters and campaigners, it can raise substantial funds if membership and small contributions are numerous enough, it can represent otherwise dispersed interests between elections, and it can compensate to some extent for the cultural hegemony of the most influential groups and institutions, advancing its own views and symbols and shielding the followers against the dominant influences.

This compensation for the impact of social and economic inequality requires, however, that these organizations are relatively autonomous from dominant groups and responsive to their constituencies. Responsiveness may be endangered by oligarchic tendencies that are hard to control. Furthermore, dominant interests quite often seek to protect themselves by sponsoring sympathetic organizations and parties with broad appeal. Many voluntary associations rely on private (though often tax supported) funding from a relatively small number of wealthy patrons. A simple measure of participation in civil society, then, is not sufficient to gauge the compensatory potential of collective organization. What is decisive is that relatively autonomous organizations protect otherwise disadvantaged interests.

The importance of autonomous collective organization in parties and politically relevant voluntary associations for leveling political inequality can hardly be exaggerated. If the economic and political power of concentrated wealth, the cultural hegemony of a limited set of elites, and the decision-making power of an imperfectly controlled state apparatus are the most important factors underlying political inequality, strong and autonomous organization of subordinate interests is the most important counterbalancing factor. It offers political competition to the influence of wealth and high status and limits the relative autonomy of the state vis-à-vis subordinate interests. It gives subordinate groups some protection against hegemonic influence by offering alternative views and orientations. And it can strengthen the universalistic norms of academic, social, and political analysis as well as of news reporting by sponsoring research informed by the concerns and interests of subordinate groups.

The conditions favoring or hindering the development of such forms of collective organization are closely related to the conditions shaping political participation generally. Prominent among them are mutually reinforcing processes in the political sphere: The prospect of political success stimulates political participation, while its absence stifles it; this seems to be a major reason why, in many countries, politically oriented participation is more prevalent among the middle classes than among the working class or the poor. Gaining or losing political influence on decisions about legislation and implementation can in turn favor or hinder collective organization.

In countries whose institutional setting has been shaped by strong unions, strong parties of the left, and a significant participation of these parties in government, class and status differences in social and political participation are much reduced or eliminated. This claim has empirical support, and on simple reflection it makes good theoretical sense. It is no real surprise that the sustained impact of successful self-organization of subordinate interests and of the representation of these interests in the governance of societies diminishes the social, economic, and cultural factors maintaining political inequality and leads to a leveling in the social and political participation of different socioeconomic strata. By contrast, the disappointing performance of recently democratized political systems in Latin America seems largely due to a renewed weakness in the organized representation of subordinate interests.

These interactions between collective organization and political success can create stable paths of advance or decline of equality in a political system, once critical turning points are passed. Therefore, the chances of a political self-organization of subordinate interests and its potential effects on political equality may well be dependent on historical conditions that are not easily changed.

Democratic Equality Versus Social and Economic Power

Political equality is a critical dimension of the quality of any system of democratic rule. It stands in tension with the structure of social, economic, and cultural inequality. Even democracy as minimally conceived is only possible if political decisions are to some extent separated from the system of class, status, and power. But since structured inequality can never be entirely eradicated and political decision making can never be fully emancipated from the inequality in power resources, democratic equality is a goal that can only be approximated at a considerable distance. At the same time, political equality is a value that is grounded in many practices and commitments common in democracies. The principle of an equal vote is only one but not the least among these.

The tension between democratic equality and the impact of differential economic, cultural, and social power is not a constant across modern democracies. There exists a great deal of variation, which ranges from democracies that remain extremely shallow in their egalitarian substance to realistic, if still limited approximations to democratic equality that are built on a significant empowerment of socially, economically, and culturally subordinate groups. To deepen democracy in the direction of greater political equality requires strong and sustained policies promoting social and economic equality.

Dietrich Rueschemeyer
Brown University
Providence, Rhode Island, United States

See also Civil Society; Class, Social; Democracy, Quality; Hegemony; Inequality, Economic; Social Stratification; Welfare State

Further Readings

Beramendi, P., & Anderson, C. J. (2008). *Democracy, inequality, and representation: A comparative perspective*. New York: Russell Sage Foundation.

Citizens United v. Federal Election Commission, 558 U.S. ___ (2010).

Dahl, R. A. (2006). *On political equality*. New Haven, CT: Yale University Press.

Huber, E., Rueschemeyer, D., & Stephens, J. D. (1997). The paradoxes of contemporary democracy: Formal, participatory, and social dimensions. *Comparative Politics, 29,* 323–342.

Pierson, P. (2001). The prospects of democratic control in an age of big government. In A. M. Meltzer, J. Weinberger, & M. R. Zinman (Eds.), *Politics at the turn of the century* (pp. 140–161). Lanham, MD: Rowman & Littlefield.

Rueschemeyer, D. (2005). Addressing inequality. In L. Diamond & L. Morlino (Eds.), *Assessing the quality of democracy* (pp. 47–61). Baltimore: Johns Hopkins University Press.

Rueschemeyer, D., Stephens, J. D., & Stephens, E. H. (1992). *Capitalist development and democracy.* Chicago: University of Chicago Press.

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics.* Cambridge, MA: Harvard University Press.

EQUILIBRIUM

See Game Theory

ETHICS

Ethics is the most general term used to characterize the different conceptions of the good as opposed to the evil and the just as opposed to the unjust. These abstract conceptions vary from one school of thought to the other. Kantianism favors an approach that focuses on principles, insisting on the role of duty and the universality of moral law. It builds a liberal ethics stressing the necessity of a public debate on the universality of core moral and political choices. In its Benthamian version, consequentialism and utilitarianism plead in favor of the maximization of pleasure and the minimization of pain, while John Stuart Mill argues in favor of the maximization of happiness and the minimization of unhappiness. Aristotelian ethics focuses on the importance of virtues that characterize the citizen of the polis. When ethics is contextualized, it is most often referred to as morality and designates a less abstract phenomenon. The different sets of values to which the good and the just are associated are then rooted in historical and social preferences. Finally, ethics is also used as a synonym of morals that characterize the domain of habits that includes ethos and mores. The etymological root of the word *ethics*—*ethos*—already includes this double meaning: ethics and habits. This threefold dimension of this notion—ethics, morality, and ethos—is of great importance for political scientists. It parallels the division between different approaches used in the discipline: normative accounts, causal explanatory analyses of social phenomena or correlation analyses, and interpretative analyses.

Political science echoes a strong interest for ethics and its related themes that has prevailed in the humanities since the early 1990s. However, the degree to which ethics is accepted as a legitimate theme of research differs from one subfield to the other. Some political scientists are still reluctant to use this terminology. It is more common for political theorists to do so because at least some of them engage in debates on normative issues. Political sociologists are more skeptical about using ethics as a concept. Since they usually maintain a position of axiological neutrality, they are more inclined to study the functional uses of morality. In comparative politics, the use of ethics as a terminology and as an abstract and absolute concept can also be a controversial issue, because ethics, some political scientists argue, has different cultural meanings. For similar reasons, ethics is also a controversial question in the field of international relations; it has, however, attracted significant attention and has become more and more popular in the literature over the past 2 decades.

The social sciences have been primarily interested in explaining and understanding ethics as a social and political phenomenon. Political scientists have drawn their inspiration from the works of major authors who have studied the genesis of ethics and its functions. They have used this knowledge to elaborate a new thinking on contemporary issues in which norms and values play a significant role. This has created a debate that is of great significance for all the branches of political science and has had a deep epistemological resonance. It also unveils tensions and quarrels between social sciences and normative theories, primarily those proposed in philosophy, that need to be elaborated. This development should encourage political scientists to engage in an interdisciplinarity debate and be open to views that lie at the close borders of their field.

The Social History of Ethics

There is a long-standing debate on the historical origins of ethics and the mechanisms that have led to the emergence and the development of ethics, morality, or morals. The work of Émile Durkheim serves as a reference for those who wish to ground morality in the community. Morality is rooted in ancient religious and ethnic communities, and religion serves as a pillar of moral values. Durkheim and his followers see morality as a collective phenomenon that includes looser conceptions of morality, such as an ethos, around which the members of the group coalesce. To Paul Fauconnet, one of Durkheim's disciples, who has made a significant contribution to the sociological study of responsibility, moral and legal sanctions are the cement of society.

Ethnic groups and the political communities such as nation-states produce values and moral codes, creating a web of meaning that has deep political consequences. Republicanism is a form of secular religion that brings social cohesion to wide collective entities that otherwise face the risk of anomy. As a collective, the French République is nourished by this sense of belonging to a national community. The French nation-state is sustained by the efforts of those who strongly believe in republican values. This community has its equivalents of cardinals and priests: high-ranking civil servants who serve as role models, who demonstrate a commitment to the well-being of the community and to strong ideals of freedom, justice, and brotherhood. The comparison with the United States is indeed very illuminating. The First Amendment to the U.S. Constitution prohibits Congress from establishing religion and from preventing the free exercise of religion, a prohibition that Thomas Jefferson described as "building a wall of separation" between church and state. Various U.S. Supreme Court decisions have provided guidelines for what the establishment and free exercise clauses prohibit. Thus, unlike in some European countries, there is no official (i.e., state) religion in the United States.

Ethics can be a powerful driving force for the development of societies. The development of capitalism relies, according to the German sociologist Max Weber, on puritan values that were predominant in America since its origins. An individualistic approach best captures the mechanisms that prevail in the emergence and then the development of capitalism. According to Weber, individuals forge their own capitalistic ethos grounded on the belief of predestination. The puritan has to show evidence of his or her worldly success to dispel any doubt about his or her salvation by God. Social scientists such as Peter Berger and David Martin have been tempted by this mode of analysis when explaining the role of Protestantism—and its new churches—in Latin America or in countries such as South Korea, where the growth of capitalism has paralleled the spread of this new creed. This phenomenon is explained in terms of correlation.

In his work on the court society, Norbert Elias has drawn our attention to another aspect of the development of mores. Individuals are part of a "civilizational process" that finds its origins in the postmedieval European court society and that leads them to greater self-restraint. This psychological phenomenon is nurtured by a historical dynamic that has important social consequences for the functioning of social organizations and state bureaucracies. Of course, several major events in the history of modern society—and most particularly World War II—serve as strong counterexamples that weaken the theory. However, the work of Elias is an interesting contribution to the study of progress and the euphemization of violence in contemporary democracies, for example, when studying the attitudes toward casualties in war, at a time when the number of casualties is changing more dramatically. Another dynamic that is of great importance in the study of ethics is the moralization of societies and the measure of moral progress.

Michel Foucault has brought to light another aspect of morality. Although the term itself is controversial when it is used in reference to Foucault's thought, morality is part of the first phase of his work, and individual ethics is at the center of his study of sexuality. Foucault explains in genealogical terms the role of utilitarianism in power relations in 18th- and 19th-century England. Utilitarianism is used when considering an architectural project for building a prison. Utilitarianism is also at the center of a critique of political power addressed by the agents of the market; Foucault refers to this phenomenon as the "critique of governmental reason" (*critique de la raison gouvernementale*). Utility becomes the cornerstone of their critique of political and particularly state decision. Political leaders and politics as a system are still being challenged by market-oriented agents and or utilitarians who frame their discourse in similar terms. In the second part of his work, Foucault discusses the modes of sexual identities and relationships in ancient Greece and highlights one of its principal aspects: the "care of the self" (*le souci de soi*). This self-knowledge reflects the virtues of the Greek citizen; it also has a political dimension to the extent that these personal traits characterize those who belong to the polis.

These social science approaches all provide an explanation of the emergence and development of ethics. As in the case of Weber, Elias, and Foucault, they also favor a study of the understanding of ethics as a language or as an element of discourse. Explaining and understanding are also intermingled with normative views. Durkheim's philosophical background is Kantianism. Weber makes a distinction between the scholar's principled approach to ethics—an ethics of conviction—and a consequentialist framework that belongs to the sphere of politics. Foucault gives an illuminating interpretation of utilitarianism and also puts into perspective the role of virtue that lies at the core of Greek philosophy and especially Aristotelian ethics. Elias also gives an example of the use of a virtue, restraint, as a trait of moral character rooted in the psyche.

The Tensions Between Ethics and Politics

Those who successfully promote ethical ideas often have access to power and to material resources. This fact has led some major philosophers to see ethics as an ideology used to subordinate the weak. The disillusionment expressed in this ontological and political vision is shared by many political scientists. Thomas Hobbes argues in favor of the use of religion as a source of legitimacy for the Leviathan that exercises his control over the church. Marxism points at the predominant role of economic and political power over ideas. Political scientists have mostly focused on one social dimension of ethics and its instrumentalization of ethics by the rulers.

Realists and neo-Marxists consider that material power prevails and that ideas play a political role when they are not an obstacle to the pursuit and maximization of the power of social and political elites. The founding fathers of classical realism have argued along those lines. Neo-Marxist approaches or theoretical frameworks inspired by the work of Pierre Bourdieu also tend to discard the role of ethics in the fields of politics and economics. When they refer to ethics, rulers use this discourse of legitimation to pursue their ends; they only play lip service to ethics and only when it does not constrain their action. On the contrary, it is a resource used to maximize their power. When discussing the role of lawyers or moral entrepreneurs, Bourdieu's followers often refer to a "market of virtue," widely dominated by those who have the resources to set the rules of the game.

However, there can be other more interesting and nuanced approaches to this question. Depending on the historical and cultural context, morality and the legitimacy derived from its sources can be a powerful resource. This, however, does not necessarily imply the structural subordination of ethical agents to power. Religious actors can be crucial players that engage in bargaining relations with princes. Indeed, over its history, the Papacy has often been courted by leaders willing to be endorsed by the Church. In contemporary politics, the role of religious groups has been highlighted when they use their bargaining powers in democratic elections.

Even classical realism is more nuanced than the general materialistic approach that is generally attributed to it, in which ethics is subordinated to power. Indeed, at the individual level, Hans Morgenthau points out the virtues of political leaders, and in doing so he rehabilitates the role of ethics in politics. Wisdom as a moral trait is an important resource for the prince or political head. The art of politics lies on some moral capacities and virtues in an Aristotelian sense.

Interests and Ideas

The interaction between interests and ideas—which include moral visions or beliefs—is one of the central conceptual dimensions of the question of moral authority. When interests determine the content of ideas, material power has the capacity to dictate the moral message of moral entrepreneurs. However, as Weber has brought to light, ideas do matter. Moral agents and the moral visions that might emanate from them can in specific contexts orient the definition of the interests of and, therefore, the decisions made by rulers.

Constructivists describe, interpret, and explain the trajectories of certain norms. Contingent visions of morality can lie at the origins of the norm trajectory as Martha Finnemore, Nina Tannenwald, and Richard Price indicate in their work on the cultures of national security. This genealogical analysis highlights the social origins of certain prohibitions. A taboo can be a source of inaction, as in the case of the nuclear weapons not used since World War II; such a taboo also parallels other abstract visions of the ethics of nuclear weapons that set a legal and ethical framework banning their use.

There are numerous critiques to this approach. Critical sociologists and realists first consider that liberal theory and constructivism overestimate the role of ethics in politics and the power of norms and moral entrepreneurs. Critical sociologists also strongly criticize liberal theorists and constructivists for their political agenda. These critics argue that liberals and constructivists are too complacent about the relationship between states and power elites and thus reproduce their own discourse of legitimation, whereas the task of the social scientist would be to deconstruct this discourse. Their epistemological flaws are said to reflect their political biases and their naïveté.

There are more convincing critiques that can be addressed to constructivist analysis of morality. The constructivist focus on moral progress can induce a bias in their vision of international relations. There is a need to differentiate between the source of the emergence of a norm—a taboo, for instance, or a natural instinct of the just and the good—and the consequence of the application or nonapplication of the norm. There is also a need to take into account the indirect consequences of the implementation of a norm. A norm might have some indirect negative consequences that are sometimes unforeseeable or remain unforeseen; thus, such a norm may not be considered as a cause of moral progress.

Universalism and Relativism

When political scientists study ethics, willingly or not, they engage in an epistemological debate with normativists, primarily philosophers. Indeed, several key epistemological issues have recently emerged and have created lines of debate and front lines in the subfields of political science. There have been some interesting discussions that all testify to the necessity of interdisciplinarity in the social sciences and in the humanities.

The question of universalism versus relativism is one of the most common issues that has created tensions both within the political science community and also in the discussions between social scientists and philosophers, as, for example, in the debates between Clifford Geertz and Richard Rorty on relativism, antirelativism, and anti-antirelativism. The position that power determines what is just and good, advocated by some political scientists, implies a strong relativist stance. The debate about democracy and democratization is an interesting case in this regard. Liberal political theorists implicitly engage in a normative debate about the moral superiority of democracies over other types of regimes. Liberal theorists argue that respecting the rule of law and ensuring pluralism are duties that a political regime must fulfill. Consensus does not prevail. Political sociologists or anthropologists would make the case that societies should have a certain degree of autonomy, with each following its own path. They therefore de facto create their own norms, values, and ethos. In the long run, do societies tend to be democratic? In the 1970s, developmental studies argued that there is a political trajectory of development leading to democracy. This mode of explanation has strong normative implications and has been severely criticized for being Western-centric, paternalistic, and hegemonic.

Are rules of justice universal? If there is such a universalism, to what extent ought it be promoted or imposed by coercive measures? This question is of crucial importance for the various branches of political science. Political sociologists argue about the reasons that motivate the imposition of moral norms and about their consequences. They also discuss the universality of norms and their social and historical contextualization. Philosophers too are deeply concerned with these issues. The theory of justice presented by philosopher John Rawls has been one of the major contributions to political theory over the past decades. Rawls's vision of justice as fairness and his concept of the veil of ignorance have greatly influenced the literature in this field and also points at other questions that go beyond political theory. Indeed, Rawls's theory of justice was elaborated in a context where individuals gather in one closed political community. How is this question to be elaborated at the international level? Would a Rawlsian approach favor the imposition of principles of justice on nonliberal states? His views on this question are nuanced. Rawls argues in *The Law of Peoples* that although liberal regimes are superior to nonliberal ones, tolerance should prevail so long as a regime does not violate human rights; however, international intervention is justified when human rights are violated.

The tendency toward an increasing social and cultural interdependence is a strong incentive to develop a new thinking in this area. "Thick" and "thin" are the two categories that Michael Walzer uses to distinguish a strong, abstract, and a priori normative judgment from one in which flexibility prevails and where moral judgment is more contextualized. In this second case, there would be a core set of principles that should never be violated; the avoidance of inflicting unnecessary pain could be such a principle. Different societies may have different interpretations of the same principle, and their members would be free to preserve their cultural identity. The core principles of a democracy, such as public liberties and the right not to be discriminated against, are inalienable, whereas moral and religious precepts that locally regulate marriage or education would be untouched as far as they do not violate core principles such as physical integrity or respect for human dignity.

Ethics as a Discussion

This phenomenon of ethics as a discussion is best captured by pragmatist philosophers. In the long-standing tradition of American philosophy that goes back at least to John Dewey, authors such as Richard Rorty argue in favor of a "relativism of justification." Ethics is thought of as a forum for debate, a public space where the most plausible, morally correct, and politically appropriate justification prevails. This implies the use of power—material power and or knowledge—to achieve this end. Ethics should also be scrutinized by the wider audience that includes other Western democracies. Therefore, ethics is a dialectical mechanism of justification. It is a justification used by people and institutions to rally the greatest number of people to their cause and to anticipate the critique they would face ex ante and ex post. It is also a weapon used by competitors who wish to block or modify the decisions taken by the majority. Finally, ethics is the output of such a dialectic. Its realm is also pluridimensional; it can take place at the local, national, international (the arena of interstate relations), transnational (the realm of transnational actors), or global levels (where states interact with transnational actors).

Ariel Colonomos
CERI-Sciences Po/CNRS
Paris, France

See also Critical Theory; Greek Philosophy; Human Rights, Comparative Perspectives; Human Rights in International Relations; Kant, Immanuel; Liberalism; Multiculturalism; Normative Political Theory

Further Readings

Durkheim, É. (1925). *Les formes élémentaires de la vie religieuse: le système totémique en Australie* [The elementary forms of the religious life: The totemic system in Australia]. Paris: Alcan.
Elias, N. (1978). *The civilizing process*. New York: Urizen.
Foucault, M. (1980). *History of sexuality*. New York: Vintage Books.
Foucault, M. (2004). Leçon du 10 Janvier 1979 [Lesson of January 10, 1979]. In *Naissance de la biopolitique Cours au Collège de France 1978–1979* [Birth of the biopolitics course to the high school of France 1978–1979] (pp. 3–28). Paris: Hautes Études Gallimard Seuil.
Katzenstein, P. (Ed.). (1996). *The culture of national security norms identity in world politics*. New York: Columbia University Press.
Morgenthau, H. (1948). *Politics among nations: The struggle for power and peace*. New York: Knopf.
Rawls, J. (1999). *The law of peoples*. Cambridge, MA: Harvard University Press.
Rorty, R. (1986). On ethnocentrism: A reply to Clifford Geertz. *Michigan Quarterly Review, 25*, 525–534.
Walzer, M. (1994). *Thick and thin moral argument at home and abroad*. Notre Dame, IN: University of Notre Dame Press.
Weber, M. (1930). *The protestant ethic and the spirit of capitalism*. New York: Scribner.

ETHNICITY

Understanding ethnicity is a challenge for political science. How is the good polity to deal with cultural differences: Should it insist that each individual be treated "equally," or should it take note of cultural variation and ethnic diversity? These questions pose challenges for normative theorists. Equally compelling are the challenges posed for students of development, a subfield in which the study of ethnicity looms large. Although there are many things that "development" can denote, this discussion limits it to two: economic prosperity and political security. Ethnicity plays an ambivalent role with respect to each, it would appear, promoting private accumulation but inhibiting public investment and multiplying opportunities for peaceful bargaining while also provoking bloodshed and violence. Even within a discussion so narrowly framed, the relationships between ethnicity and development are complex. This entry discusses these relationships from a political economy perspective.

Prosperity

The formation of capital underpins the growth of prosperity. When people form capital, they refrain from consumption and set a portion of their earnings aside; by investing that portion today, they seek to secure higher levels of consumption tomorrow. When the gains from the investment outweigh the losses from the initial sacrifice, then the result is an increase in welfare for the individual and economic growth for the society.

When we speak of investment, we are speaking of capital formation. Families play a major role in capital formation and, in particular, in the formation of human capital. They do so by structuring relationships between generations such that the older generation, rather than consuming all its income, instead devotes a portion of its resources to the younger, who in turn will prosper and, while doing so, offer support to those who had previously invested in them.

This intergenerational flow of funds underpins two forms of capital formation, and ethnic groups play a major role in both. One is education, another migration. The two are closely related, of course, as by educating their children, parents seek to prepare them to compete in urban labor markets.

To illustrate, consider Victor Uchendu, a noted anthropologist, who writes of his childhood in eastern Nigeria. Uchendu recalls the pride he felt as a child as his academic abilities became apparent. He writes, too, of the sacrifices his parents made to pay for his schooling and of the contributions collected by his community to send him for further education abroad. Each step of the way—from village school room, to secondary school, and thence to the university—opened up a broader range of opportunities. Uchendu also writes of the burden of his success; for those who made it possible then looked to him for leadership, hoping that he, as an "Ibo son abroad," would facilitate their own transition from villager to a member of the global community.

The pressures placed on Uchendu find their parallel in other upwardly mobile communities. The Kikuyu of Kenya, the Chinese of Malaysia, the Jews in Europe, and the Bengali in India: These and other ethnic groups have pursued a strategy of collective improvement, based on the transfer of resources between generations and the formation of human capital.

Already suggested is the close link between education and migration: Investments in the former are often made to ensure success in the latter, as the skills acquired in school translate into better jobs in urban labor markets. Migration, it has been found, like education, represents a strategy formulated by families rather than a decision made by individuals. With globalization, moreover, the urban destination need not be an adjacent city or the national capital, but rather Houston, London, or Paris. Reflecting the investment-like character of the process, the Northern flow of people is mirrored by a Southern flow of remittances. But why do these assume an ethnic character?

One reason is that as educated persons from the countryside flow into labor markets in towns, their skills confer a competitive advantage. Locals often respond by demanding that places in the classroom be reserved for them and that the government guarantee them a "fair share" of the jobs. As "sons of the soil," they demand preferential treatment (Myron Weiner, 1978). Such responses characterized those of the Yoruba tribe in the face of Ibo immigration, the Malays in response to the Chinese, and the local population of Assam when challenged by the Bengali—the list is long.

Less apparent but no less powerful is the role of ethnicity in strengthening the expectations that shape the relationships between generations. It is the adults who sacrifice and invest and the youths who benefit and repay. The sacrifice is immediate; the repayment is necessarily delayed. While both parents and offspring should benefit from the investment, it is the parents who bear the risk; should they fear the youths' defection, the pact between generations will break down. The role of the ethnic group is to reduce such fears.

One way is by celebrating the family bonds. The groups celebrate the blood their members share and the ties of kinship that bind them. They honor the contributions of previous generations and the debt that is owed. Inculcating such values in the young helps assure the old that the youths will fulfill the intergenerational contract. In addition, in some settings, ethnic groups possess power. If a group possesses a homeland, the elders are likely to occupy political and religious offices. Being in a position to sanction youths who abscond, the elders may then have reason to believe that few will actually do so (Robert Bates & Irene Yackolev, 2002).

Public Goods

While ethnic groups may thus contribute to prosperity by promoting the formation of human capital, they also appear to impede it. They do so when the diversity of preferences among ethnic groups reduces a community's ability to supply public goods or to share in the costs of their provision. A transport system, the supply of electricity or water, or a harbor or airport, each requires large public investments. And the costs per capita—that is, the tax burden—will be lower the larger the size of the community that provides them. But as the size of the population increases, so too will the heterogeneity of preferences. One group might prefer that a road run through its territory rather than elsewhere; another may prefer that the airport not be located in "its back yard." The ability to apportion the costs and benefits declines as the number and intensity of such disagreements rises, making it more difficult for multiethnic societies to contract for the supply of public services. Ethnic diversity is an important source of such heterogeneity, and a great deal of research has been devoted to its impact. The relationship between diversity and public goods provision is "one of the most powerful hypotheses in political economy" (Abhijit Banerjee, Lakshmi Iyer, & Rohini Somanathan, 2005, p. 629).

James Habyarimana, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein (2009) deepen our understanding of the relationship between ethnicity and public goods provision. Using experimental methods in a field setting, they probe the microlevel linkages that underlie the macrolevel patterns described in this entry. While Bates believes that Habyarimana et al. (2009) fail sufficiently to replicate conditions of external threat—conditions commonly associated with ethnic politics—they significantly advance our understanding of the mechanisms that impart causal force to ethnic relations.

Ethnicity is thus Janus-like. On the one hand, ethnic groups provide governance structures that promote private investments in education and migration. For individual families, then, they strengthen the prospects for prosperity; and for societies, they strengthen the forces of economic modernization. On the other hand, where there is ethnic diversity, this tends to weaken the ability of the state to supply public goods. Insofar as these services are themselves valuable or underpin private production, ethnic diversity lowers the level of diversity and the level of development.

Conflict

In addition to raising the costs of providing public goods, ethnic diversity—it is held—poses an additional threat: the prospect of conflict. While intuitively appealing, there is little systematic evidence to support this claim. Rather, once again, a more nuanced view emerges from the literature.

The equation of ethnic diversity and conflict stems from an obvious reality: In much of the developing world, where there is conflict, ethnic groups are likely to be involved. But, as noted by Bates and Yackolev (2002) and James Fearon and David Laitin (1996), while ethnic diversity may characterize developing polities, conflict remains a rare event: Even where there is a high level of ethnic diversity, there is often little fighting. The implication is clear: As a rule, ethnic groups cohabit peacefully. And, indeed, cross-national regressions find *no* systematic relationship between measures of ethnic diversity and the likelihood of civil conflict.

In general, ethnic politics in the developing world resembles that in the major metropolitan areas in the developed world. Politicians draw on ethnic blocs of voters, extract enough resources to offer enough patronage to retain their positions, and bargain with others similarly positioned.

Having set aside the conventional wisdom and clarified the basic reality, the fact remains that ethnic groups do fight each other and that ethnic clashes number among the most deadly forms of civil conflict. It is important to determine, then, the conditions under which ethnic politics turn violent.

They do so, it would appear, when a group faces the prospect of political exclusion. When in 1962 politicians from the northern region of Nigeria allied with those in the western region, for example, the Ibo in the eastern region "stood alone," in the words of their leading newspaper; their isolation marked an important step toward the civil war that resulted from their attempted succession. More data come from India, where Steven Wilkinson (2004) finds a systematic relationship between the threat of exclusion and political violence in municipal elections. In constituencies that offer multiple ways of forming majority coalitions, he finds, politicians face incentives to bargain in order to construct a majority coalition, and the composition of that coalition may alter between elections; when the constituency is evenly divided and the divisions likely to be long-standing, however, political leaders may seek to provoke incidents that will politicize and rally their backers. The expected value of the risky outcome from violence may exceed the losses certain to follow from the exclusion from power.

Some have noted and stressed the spatial as well as the size distribution of ethnic groups; others have emphasized the extent to which they differ from each other, not only culturally and linguistically but also in their choice of religion and their incomes. Such features are difficult to summarize numerically, which hinders efforts to test their significance for ethnic conflict. Thus far, only indices of ethnic polarization have been computed, and they have in fact been found to be positively related to political conflict.

Politicians may also provoke ethnic violence when the loss of office would leave them vulnerable to reprisals. This is particularly true, of course, of those who have themselves used violence while in office. Examples can be taken from the period of redemocratization in Africa (roughly between 1989 and 1994). Some, such as Samuel Doe of Nigeria, purposively attacked the traditional rivals of his own ethnic group so that the latter would be unable to defect from his coalition and thus lose the protection he, as president, could offer. A similar situation was seen in Rwanda when the Hutu elite appeared to have provoked retaliatory attacks by Tutsi in an effort to obviate the formation of a cross-ethnic coalition that would have threatened their rule. Statistical support is offered by Paul Collier and Robert Bates (2008), who find that when rulers have used violence to consolidate their power, ethnic diversity increases the likelihood of conflict. While there is thus no universal relationship between ethnic diversity and political conflict, there appear to be particular conditions that render ethnic violence significantly more likely.

Conclusion

Recent work has thus deepened our understanding of the relationship between ethnicity and development. In light of this work, the sweeping claims of earlier scholars (David Apter, 1963) and contemporary commentators (Robert Kaplan, 1994) give way to a more nuanced assessment. We now know that ethnicity can both help and hinder the process of development and also help one to have a clear idea of the conditions under which it will do the one or the other.

Robert H. Bates
Harvard University
Cambridge, Massachusetts, United States

See also Clientelism; Conflicts; Discrimination; Identity, Social and Political

Further Readings

Alesina, A., & Spolare, E. (2003). *The size of nations.* Cambridge: MIT Press.

Banerjee, A., Iyer, L., & Somanathan, R. (2005). History, social divisions, and public goods in rural India. *Journal of the European Economic Association, 3*(April–May), 639–647.

Bates, R. H., & Yackolev, I. (2002). Ethnicity, capital formation, and conflict: Evidence from Africa. In C. Grootaert & T. van Bastelaer (Eds.), *The role of social capital in development* (pp. 310–340). New York: Cambridge University Press.

Collier, P., & Bates, R. (2008). Endogenizing syndromes. In B. J. Ndulu, S. O'Connell, R. H. Bates, P. Collier, & C. Saludo (Eds.), *The political economy of economic growth in Africa: 1960–2000* (pp. 391–420). Cambridge, UK: Cambridge University Press.

Easterly, W., & Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics, 112*(4), 1203–1250.

Fearon, J., & Laitin, D. (1996). Explaining interethnic cooperation. *American Political Science Review, 90*(4), 715–735.

Fearon, J., & Laitin, D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review, 97*(1), 75–90.

Habyarimana, J., Humphreys, M., Posner, D. N., & Weinstein, J. M. (2009). *Coethnicity: Diversity and the dilemmas of collective action*. New York, Russell Sage Foundation.

Hershfield, A. F. (1969). *Ibo sons abroad: A window on the world*. Montreal, QC: Annual Meetings of the African Studies Association.

Kaplan, R. D. (1994). The coming anarchy. *Atlantic Monthly, 273*(2), 44–76.

Strauss, S. (2006). *The order of genocide: Race, power, and war in Rwanda.* Ithaca, NY: Cornell University Press.

Uchendu, V. C. (1965). *The Igbo of Southeast Nigeria*. New York: Holt, Rinehart and Winston.
Weiner, M. (1978). *Sons of the soil: Migration and ethnic conflict in India*. Princeton, NJ: Princeton University Press.
Wilkinson, S. I. (2004). *Votes and violence*. New York, Cambridge University Press.

Ethnographic Methods

Ethnography is a methodology with more than 100 years of history. It originated in the Western world as a particular form of knowledge about distant cultures (typically non-Western ones) that were previously largely unknown to the outside world, having had only fleeting contacts or brief conversations with some persons. After having entered several disciplines of the social sciences (anthropology, sociology, and psychology), in recent times ethnography has become an important tool in political science as well (where it takes the name of *political ethnography*), as a part of interpretive methodologies and research. This entry discusses the origins and definitions of this method, its strengths and weaknesses, and its increasing relevance for political science.

Competing Definitions of Ethnography

Defining a term is always difficult because there are as many definitions as there are different points of view. Paul Atkinson and Martyn Hammersley observe that the definition of the term *ethnography* has been subject to controversy. For some scholars, it refers to a philosophical paradigm to which one makes a total commitment; for others it designates a method that one uses as and when appropriate. But the controversy extends further. Since the 1980s, the meaning of ethnography has been expanded to such an extent that it encompasses forms of research that are extremely diverse from a methodological point of view. Everything is now ethnography: from life stories to analysis of letters and questionnaires, from autobiography to narrative analysis, and from action research to performance to field research lasting from a few days to several years. For example, leading scholars such as James Lull and David Morley have pointed out that what passes as ethnography in cultural studies fails to fulfill the fundamental requirements for data collection and reporting typical of most anthropological and sociological ethnographic research. Ethnography has become an abused buzzword and has been diluted into a multitude of sometimes contrasting and contradictory meanings, becoming synonymous with qualitative studies.

Amid this increasing polysemy of meanings, there are at least three terms that are related to ethnography: participant observation, fieldwork, and case study. However, these should not be confounded. The term *case study* denotes research on a system bounded in space and time and embedded in a particular physical and sociocultural context. Research is conducted using diverse methodologies, methods, and data sources, such as participant observation, interviews, audiovisual materials, documents, and so on. The term *fieldwork* emphasizes the continuous presence of the researcher in the field, as opposed to the grab-it-and-run methodologies, such as surveys, in-depth interviews, or analysis of documents and recordings. In this case, too, diverse methodologies and methods may be used. Finally, *participant observation* is a distinct research strategy. Participant observation and fieldwork treat observation mostly as a mere technique, whereas the term *ethnography* stresses the theoretical basis of such work stemming from a particular history and tradition.

An Updated Definition

The stretching of the term *ethnography* has emptied it of its original meaning. Ethnography was born as a technique based on direct observation. By contrast, interviews and surveys are mainly based on listening and asking questions. Of course, it is also essential in ethnography to listen to the conversations of the actors "on stage," read the documents produced by the organization under study, ask people questions, and so on. Yet what most distinguishes ethnography from other methodologies is the role of the "protagonist" assigned to observation. Ethnographic methodology comprises two research strategies: *nonparticipant* observation and *participant* observation. In the former case, the researcher observes the subjects "from a distance" without interacting with them. Those who use this strategy are uninterested in investigating the symbolic sphere, and they make sure not to interfere with the subjects' actions so as not to influence their behavior. Of course, there are many intermediate situations between the two extremes of participant and nonparticipant observation.

Participant observation has the following characteristics: (a) the researcher establishes a direct relationship with the social actors, (b) staying in their natural environment, (c) with the purpose of observing and describing their behavior, (d) by interacting with them and participating in their everyday ceremonials and rituals, and (e) learning their code (or at least parts of it) to understand the meaning of their actions.

A Historical Sketch

The birth of ethnographic methodology is commonly dated to the period between the late 19th and early 20th centuries. It was developed in ethnology, a discipline that in the first half of the 1800s split from traditional anthropology, which was then dominated by physical and biological concepts. Ethnology was more concerned with studying peoples (through comparison of their material artifacts) and their cultures and classifying their salient features. Before the advent of ethnographic methodology, ethnologists did not collect information by means of direct observation; instead, they examined statistics, the archives of government offices and missions, documentation centers, accounts of journeys, archaeological finds, and native manufactures or objects furnished by collectors of exotic art, or they conversed with travelers, missionaries, and explorers. These anthropologists considered the members of native peoples to be "primitives": They were savages to be educated, and they could not be used as direct informants because they could not be trusted to furnish objective information.

Ethnographic methodology did not suddenly erupt in anthropology; rather, it developed gradually through the work of various authors, among them the English anthropologist of Polish origin Bronislaw K. Malinowski (1884–1942), and the English anthropologist Alfred R. Radcliffe-Brown (1881–1955). British social anthropology assimilated the positivist intellectual climate of its time and put itself forward, according to Radcliffe-Brown (1948), as a "natural science of society" that was better able to furnish an objective description of a culture than the other methods used by anthropologists at the time. Radcliffe-Brown's polemic was directed against the then dominant speculative or "desk" anthropology, which preferred to rely on secondary sources rather than to conduct direct observation of social facts (customs, rituals, and ceremonies) in order to uncover the "laws" that govern a society.

Malinowski is commonly regarded as being the first to systematize ethnographic methodology. In his famous *Introduction to Argonauts of the Western Pacific*—the book that sets out his research conducted in the Trobriand Islands of the Melanesian archipelago off eastern New Guinea—Malinowski described the methodological principles underpinning the main goal of ethnography, which is to grasp the native's point of view and his relation to life and to realize his vision of his world. For this purpose, Malinowski lived for several years during World War I among the Kula of the Trobriand Islands. He learned their language (Kiriwinian), used natives as informants, and directly observed the social life of a village, participating in its everyday activities. Malinowski inaugurated a view "from within" that American anthropologists of the 1950s would call the "emic" perspective—as opposed to the "etic" or comparative perspective, which instead seeks to establish categories useful for the analyst but not necessarily important for the members of the culture studied.

From the 1920s onward, ethnographic methodology was incorporated into sociology—where it was adopted by researchers who mostly belonged to the Chicago School—and then into psychology and (recently) political science. It was imported from anthropology, however, for 70 years earlier the French mining engineer and later sociologist Pierre Le Play (1806–1892) had used primitive forms of participant observation when he had stayed with the working-class families that he was studying. The English philanthropist Seebohm B. Rowntree (1871–1954) also used early forms of participant observation (after 1886) for his inquiries into poverty and living conditions in the London slums.

Ethnography and Political Science

Ethnographic methods have recently also entered political science, long dominated by official statistics, surveys, and comparative methods. The entry of ethnography has been favored by two cultural and theoretical changes in the discipline: an interest in the "micro" dimensions of political phenomena and an openness to the insights of qualitative research. These two changes have been carried forward by important scholars such as Sidney Tarrow (1995) and Charles Tilly (2001), who proposed bridging qualitative and quantitative approaches by means of "triangulation" to explore the micro–macro link, considering that a purely quantitative approach fails to uncover the causal mechanisms of political processes.

As said with regard to other disciplines, the term *ethnography* has assumed a variety of meanings in political science, and it has become largely synonymous with *fieldwork*. The work of Jeremy Weinstein (2007) on political violence and civil wars and Elizabeth Jean Wood (2000) on democratic transitions in South Africa and El Salvador are examples of this kind. Drawing mainly on narrative interviews (with the addition of some observations, official statistics, and government documents), these authors have sought to reach the experiences, the subjective perspectives, and the points of view of people involved in violent actions and democratic transitions.

Weinstein attempted to uncover specific causal mechanisms (why some rebel groups decide to use indiscriminate violence against civilians) by going beyond traditional quantitative studies, which explain this phenomenon of violence through macrovariables such as income and so on. He analyzed the inner dynamics (the recruitment of members and the inner hierarchical structure) of rebel groups in Peru, Mozambique, and Uganda. He found that indiscriminate violence against civilians was committed mainly by rebel groups with external financial resources, such as those deriving from drug trafficking, foreign aid, and so on. Consequently, these groups did not need or seek civilians' involvement in and consensus on their political actions.

Wood noted that South Africa and El Salvador are very different when structural variables (economic development, race composition, etc.) are considered but quite similar when the link between the political elite and the economic elite is examined. She also focused on the bottom-up, violent mobilization of workers, pointing out that this produces an increase in costs affecting the entire production system (and, consequently, the economic elite)—strikes, vandalism, economic uncertainty, a decrease in foreign capital inflows, and so on. To resolve the situation, the economic elites push for reform of the authoritarian regime and the economic and productive system that supports this form of government. Wood pinpoints everyday political processes, the impact of local actions on national politics, and the effects of micro-events on macrophenomena.

In addition to studies based on fieldwork, now appearing in political science are real ethnographies, such as the work of Adam Ashforth (2005) on violence and democracy in South Africa. Ashforth, a White American, spent 3 years as a guest of a family in Soweto, the well-known Black township on the outskirts of Johannesburg. During this time, Ashforth realized that it would be extremely difficult to understand local politics (macro) without considering witchcraft (micro) and its role in social relations (meso). Through participation in the everyday life of the community, he acquired the conceptual categories, the constellation of meanings, and the culture of Soweto's residents. He learned that witchcraft beliefs were remedies for the uncertainty and insecurity of everyday life; that envy and jealousy produced the social conditions for the success of witchcraft; and that the latter shaped relations among individuals, social groups, and political institutions on issues such as the spread of AIDS, its social consequences, and health policies. He also discovered the effects of the building of democracy on community members, their acceptance of violence, the shape of the concept of social justice, and the affirmation of a modern democratic and liberal state.

The Added Value of Ethnography

As mentioned before, ethnographic methodology gives priority to observation as its primary source of information. This purpose is also served, in a secondary and ancillary manner, by other sources of information used by the ethnographer in the field: informal conversations, individual or group interviews, and documentary materials (diaries, letters, class essays, organizational documents, newspapers, photographs, and audiovisual aids). However, the overriding concern is always to observe actions as they are performed in concrete settings. As John

Heritage notes, if one is interested in action, the statements made by social actors during interviews cannot be treated as an appropriate substitute for the observation of actual behavior. In fact, there is an often documented gap between attitudes and behaviors, between what people say and what they do. Ethnography is therefore the privileged methodology with which to analyze and understand social and political actions in their making.

The presence of the researcher in the field enables him or her to gain better understanding of the conceptual categories of social actors, their points of view (emic), the meanings of their actions and behavior, and social and political processes. As Javier Auyero (2006) maintains, political ethnography highlights the aspects neglected by quantitative analysis, such as the impact of micropolitics on macrophenomena, the complexity of everyday life, the network of participants' meanings, their motivations, the making of political action, and the practices of politics. As Charles Tilly (2006) notes, ethnographic methods are useful for the analysis of political phenomena consisting not of macrostructures and fixed roles but of interactions among participants, families, and small groups, emancipating an inquiry from the ethnocentrism (etic) that still characterizes many explanations of political science.

Ethnography and Its Enemies

Notwithstanding the acknowledged usefulness of ethnography, however, it is still subject to several well-known stereotypes and prejudices.

Is Ethnography a Highly Subjective Method?

It is often argued that ethnography is a purely subjective method, in the sense that it is very sensitive to the researcher's attitudes and perceptions. In other words, if different researchers visit the same setting, they will see different things, and their ethnographic notes will record different aspects. Instead, a questionnaire or a discursive interview, if conducted correctly, is more likely to obtain similar replies (reliability) regardless of who the interviewer is. And yet experience shows that this idea has weak empirical foundations. For example, some time ago, Giampietro Gobo's students conducted an ethnographic study in a pub. Two groups (formed of three students each) visited the same pub a few days apart. The fact that they had chosen the same pub was absolutely coincidental, in the sense that they had not agreed on it beforehand. Nevertheless, the two groups had a specific research design: to study the rituals, ceremonials, and behaviors of consumption in pubs. They then produced a report. Gobo discovered, to his great surprise, that they had observed and discovered practically the same things. Therefore, the research design makes a greater contribution to discovery (or construction of data) than does the researcher himself or herself. Ethnography, therefore, is anything but a *highly* subjective methodology (even if subjectivity is ever present, as in all methodologies).

Behaviors Are More Consistent Than Attitudes and Opinions

What does the experience just described tell us theoretically? In other words, why did six different observers in the same pub notice practically the same things? Because what ethnography mainly observes are behaviors (rituals, routines, and ceremonials), and these are much more stable over time than are attitudes and opinions (the privileged fields of inquiry for discursive interviews and surveys), as shown by Richard La Piere's well-known experiments in the early 1930s. Those who deal with organizations know very well that altering a behavior requires more time than altering an attitude, not to mention opinions, which are sometimes so volatile that they change from one day to the next.

Can Ethnographic Research Be Replicated, Reproduced?

From this it follows that, because behaviors are temporally rather stable, the results of ethnographic research can be repeated and reproduced, with two requirements:

1. the presence of a precise research design that has guided the research and
2. no significant changes between one study and the next.

Ethnography and Generalization

A recurrent criticism made of ethnographic methods is that their results are impossible to generalize because they are based on few cases, sometimes on only one. However, there are numerous disciplines that work on a limited number of cases: for instance, paleontology, archaeology, geology, ethology, biology, astrophysics, history, genetics, anthropology, linguistics, cognitive science, and psychology (whose theories are largely based on experiments and therefore on research conducted on nonprobabilistic samples and on few cases). According to Howard Becker, these disciplines are unconcerned about their use of only a handful of cases to draw inferences and generalizations about thousands of people, animals, plants, and other objects. Moreover, science studies the individual object/phenomenon not in itself but as a member of a broader class of objects/phenomena with particular characteristics/properties (Geoff Payne & Malcolm Williams, 2005; Williams, 2000).

For these reasons, it is anything but odd to think that the results of ethnographic research cannot be generalized. As Randall Collins stated, much of the best work in sociology has been carried out using qualitative methods without statistical tests. This has been true of research areas ranging from organizational and community studies to microstudies of face-to-face interaction and macrostudies of the world system. In addition, if the focus of ethnography is on behaviors, and given that these are stable in time, then it is likely that generalizations are possible. Obviously, precise criteria must be followed in the choice of samples. Nevertheless, ethnography is not precluded from making generalizations.

Sampling Cases or Instances

It will by now be clear that the term *case* is used ambiguously in ethnographic research. In surveys and discursive interviews, the cases correspond to the number of persons interviewed (the sample) and who are usually interviewed only once. Indeed, it is rather rare for several interviews to be conducted with the same person (during a single piece of research). Then, statistical calculations and analyses of the interviews are performed on these cases.

Ethnographic research is very different. What is usually referred to as the "case" (the organization or the group studied) is in fact the setting. The cases are instead the hundreds of instances (pertaining to rituals, ceremonials, and routines) that the researcher observes or the dozens of individuals that he or she meets dozens of times during his or her presence in the field. The researcher is not interested in the organization (or the group) per se but rather in the behaviors that take place within it. Consequently, to not create confusion with the other methodologies, it would be better in ethnographic research to abandon the term *case* and replace it with that of *instance*.

The Future of Ethnography

Even if ethnography is a methodology with more than 100 years of history, it is perceived (especially in political science) as a new method. The reason why ethnography is now becoming so fashionable, probably, concerns not its inner features but the sociohistorical period in which we live. If the relationship between science and society (on how social beliefs have influenced knowledge and scientific theories and vice versa) has been widely proven and if also the relationship between technology and society (on how the birth of certain artifacts, e.g., the bicycle or the personal computer, are related to the type of society that has produced them) has been tested, it seems necessary to enlarge these perspectives by acknowledging the relationship between society and research methods—the interdependence among social beliefs and methodologies. Ethnography and society are mutually constitutive: On the one hand, ethnography (like surveys and in-depth interviewing in a previous age) requires a particular type of society for it to come into being and develop; on the other hand, these research methods strengthen the society that has produced them. We are probably entering the "observation society"—a social formation in which watching and scrutinizing are becoming the dominant cognitive modes alongside the others, such as listening, feeling, hearing, and eavesdropping, typical of the "interview society" from which survey methodology originated. This phenomenon also accounts for the increasing demand in various sectors of society—from marketing to security and from television to the fashion industry—for observation and ethnography. Ours are becoming observation societies. Amid the increasing popularity of interpretive methods, driven by epistemological approaches such as constructivism, postmodernism, feminism, and relativism, political ethnography is in a position to play a major role. The emic perspective, which aims to capture the social actors' points of view, perspectives, meanings, motivations, and emotions, has an important place in political science methods.

Giampietro Gobo
Università degli Studi di Milano
Milan, Italy

See also Case Studies; Grounded Theory; Participant Observation; Quantitative Versus Qualitative Methods; Thick Description

Further Readings

Ashforth, A. (2005). *Witchcraft, violence and democracy in South Africa*. Chicago: University of Chicago Press.

Auyero, J. (2006). Introductory note on politics under the microscope [Special issue on political ethnography I]. *Qualitative Sociology, 29,* 257–259.

Gobo, G. (2008). *Doing ethnography*. London: Sage.

Payne, G., & Williams, M. (2005). Generalization in qualitative research. *Sociology, 39*(2), 295–314.

Tarrow, S. (1995). Bridging the quantitative-qualitative divide in political science. *American Political Science Review, 89*(2), 471–474.

Tilly, C. (2006). Political ethnography as art and science. *Qualitative Sociology, 29,* 409–412.

Weinstein, J. (2007). *Inside rebellion: The politics of insurgent violence*. New York: Cambridge University Press.

Williams, M. (2000). Interpretativism and generalization. *Sociology, 34*(2), 209–224.

Wood, E. J. (2000). *Forging democracy from below: Insurgent transitions in South Africa and El Salvador*. New York: Cambridge University Press.

Europe as an International Actor

Considering Europe as an international actor continues to be a contested perspective in the study of European politics and international relations, even though scholars and decision makers increasingly see Europe as an important player within the global political system. Thus, the meaning of the phrase *Europe as an international actor* needs a few clarifications to avoid misunderstandings and disputes. *Actor* is not a controversial concept in political science. Broadly speaking, it means a unit (person, group, or organization) able to select and carry out actions aimed at influencing the subjects, institutions, and policies of the political system. Accordingly, *international actor* is the unit capable of making decisions aimed at influencing other international actors and the institutions and policies of the international political system. States as well as international organizations, nongovernmental organizations, and associations of various kinds are considered to be international actors—that is, units having attributes such as rationality and goal orientation, autonomy, and the control of resources and instruments for acting internationally. Consequently, considering Europe as an international actor means recognizing that Europe has the attributes necessary to influence by intention the actors, institutions, and policies of the international system. However, Europe is an international actor as far as the European countries agree, voluntarily and by using institutionalized mechanisms of decision, to act cohesively for influencing the international system. However, this condition applies only to the 27 countries that are members of the European Union (EU). Therefore, in this discourse, Europe is the same as the EU. In fact, the EU member states are able to make decisions on common actions toward the world system because they have the same image of the external world, agree on a cluster of common international values, share important international interests, and pursue common external goals. Furthermore, they have created policy-making offices and mechanisms in the areas of international politics and economic external relations and are determined to improve these offices and mechanisms, expand the international action capabilities of the common institutions, and make efforts for achieving better external policies. Finally, the international role and actions of the EU can count on the broad support of international actors, such as states, international organizations, and nongovernmental organizations.

A Europe of Common Values and Interests

How can one explain this common pursuit of values and interests on the part of the states belonging to the EU? Interdependence forces states to cooperate in order to produce compatible rules and policies regarding common problems such as environment degradation, transnational crime, migration, public security, and so on. This pressure is especially urgent in the case of states that, like the European ones, are in general small- and medium-sized countries placed in a rather small continental space; have an economy highly dependent on the world market; and share a democratic regime controlled by people and the media. This pressure drives governments to coordinate domestic and foreign policies in order to implement the logic of scale. Generally speaking, the larger the scale of execution of political decisions, the more appropriate and efficient will be the policy solutions of the common problems. The European governments react rather swiftly to the requisites of the politics of scale because they know well the benefits of such a strategy due to the long process of economic coordination and integration that they have passed through in the past 50 years. Thus, the European resolve to become an international actor is founded on an awareness of international interdependence and the perceived success of joint custom, trade, monetary, and other economic and sectoral policies.

However, not all analysts agree on this interpretation of the European ambition of becoming an international actor. Some of them underline the complexity and consequent slowness of the building process of the common foreign and security policy (CFSP). These analysts concede only that the EU is an international actor in the making, not an existing one. Others question the capacity of member states in building a common foreign policy and point to the national differences that make it almost impossible for the EU to solve existing policy differences viewed in the short and medium term. To other analysts, the success of an international actor depends on its economic strength as well as on the appropriateness and consistency of its strategy in global economic negotiations. These analysts hold that the EU's role as an international actor is mostly dependent on the strength of the external economic relations of the EU and is restricted to this area of international relations. Finally, some analysts claim that Europe's attempt to be an actor in international politics has been linked to Europe's ability to perform as a civilian power in the diffusion and defense of new values such as environmentalism, cultural pluralism, and human rights. These analysts point to the potentially negative consequences of current attempts to establish a common security and defense policy (ESDP) for the international prestige and leverage that the EU has achieved thanks to its role as a civilian power, advising European leaders that the best way to exert influence as an international actor is to limit themselves to exerting normative power.

This criticism notwithstanding, the convergence of the EU member states toward an important set of international values and goals is apparent in recent EU treaties. This convergence has made possible the construction of a European conception of the world system and the consequent ascription to the EU of the role of promoter of a definite set of international values and interests. In the treaties of Nice and Lisbon, the signing governments proceed from a pluralist and communitarian image of the world system and represent the EU as an actor engaged in defending values related to this image. Europe's image of the world is a pluralist one because, in addition to states, entities of a different kind—that is, individuals, peoples, and nongovernmental organizations and associations—are presented as important actors within the world system. Furthermore, the defense of new, nonstate values is considered as equal in importance to the preservation of state values. In a world encompassing individuals, organizations, peoples, and states as equals, communitarian solidarity and mutual respect among all the subjects must be promoted. States, in particular, are called on to respect the communitarian principle of mutual recognition by all the subjects and therefore to rigorously adhere and contribute to the development of international law and the principles of the United Nations (UN) Charter. In harmony with this pluralist and communitarian view, the European governments see the EU as a legitimate international actor that wants to defend different values such as peace and security, sustainable development, free and just trade, elimination of poverty, and the defense of all human rights.

Furthermore, the promotion and defense of the European identity, territorial integrity, and EU-specific values and interests are also claimed by the European leaders to be goals of European international action. In other words, the EU wants to participate in a world of peoples, individuals, and states also in order to promote its own values by all the means at its disposal. Hence, on the one hand, the EU's pluralist and communitarian image of the world is close to the conception of Europe as a civilian and normative power. On the other hand, this does not exclude a vigorous defense of European interests and specific values. Consequently, the formal image of the world contained in the treaties must be complemented by adding the position the EU leaders affirmed in the document *A Secure Europe in a Better World*, the so-called European Security Strategy, issued in 2003, as a result of the agreed decision to build the common defense policy. In this document, the EU forcefully declared a preference for effective multilateralism as the fundamental instrument of any international action aimed at fostering the development of a stronger international society, the efficient operation of international institutions, and an international order based on international law. As will be seen below, multilateralism is, indeed, the blueprint for European strategies of action in the international system.

Europe as a Security Actor

The process of convergence toward shared international interests is less definite and narrower in scope than the convergence of values. Strengthening international stability and containing conflict are interests shared by all the European states. In this view, creating conditions for the security of all the member states contributes to consolidating the common economic growth that continues to be the paramount objective of European unification. In addition, European states accept that acting together as a single international actor, especially in areas close to Europe, is more efficient and productive than acting individually; this not only provides for more resources, but neighboring countries are also less apprehensive of joint external EU actions than of the actions of individual states. Finally, convergent interests are promoted by globalization and interdependence in as much as these processes extend the international public space well beyond diplomatic and economic matters. Contemporary international relations encompass a plurality of dimensions such as the physical environment, human rights, transnational labor, and crime organizations; these cannot be dealt with efficaciously by means of independent national policies. Thus, the logic of scale is clearly perceived by all the European countries also in the area of foreign policy. Consequently, the EU governments normally agree to discuss and negotiate among themselves first in order to reach a joint position on issues of world politics. In short, having a joint European position at international conferences is viewed as the most feasible way for advancing the interests of each EU member state.

However, this pressure for interest convergence notwithstanding, the administrative and diplomatic structures of foreign policy making of the member states have not been dismantled and are far from being dismantled, for the reason that European governments want to preserve their own instruments for promoting and defending those national interests that they consider as not fully compatible with the common interests as defined by the EU itself. Consequently, in the medium to long term, EU international actions will be the output of what has been called European foreign policy. This, indeed, is the sum of three elements: (1) the common policy of external economic relations, (2) the formal dimension of the CFSP and ESDP, and (3) the national foreign policies that, at the same time, are independent from one another but under the institutional influence of the CFSP/ESDP and the views of other member states.

Europe as an International Actor in Practice

Within this framework that encompasses shared international values, a common image of the world, and the converging international interests of the European countries, this entry now discusses the presence of Europe as an international actor in the world system by reviewing Europe's strategy and actions in crucial areas of contemporary international relations, such as the politics of regions and interregional affairs, the UN, and world economic affairs.

The process of integration that sustained the development of the EU over the past 50 years and the pan-European cooperation process known as the Helsinki Process, lasting for about 30 years, from 1971 to the end of the past century, is the seminal experience that created a European preference for the so-called global approach to the solution of problems at the regional and region-to-region levels. Over the past 50 years, European countries have recognized that peaceful relations and cooperation among neighboring countries can arise from the balanced and multidimensional management of political, economic, and social interactions. On this premise, European governments constructed the principle that any cooperative program must be "global"—that is, it must encompass the political and security dimensions; the economic and financial dimensions; and the social, cultural, and human dimensions. The myth of European civil and normative power—articulated by academics and encouraged by the EU institutions and officers—is also at the root of the vision of what European foreign policy values are and what drives the EU's role as an international actor. Indeed, the global approach lies at the core of the EU grand strategy as an international actor. The late Euro-Mediterranean Partnership, now superseded by the Union for the Mediterranean, the European Neighbourhood Policy, and the region-to-region cooperation programs, such as ASEM (Asia-Europe Meeting) and the EU-Gulf Cooperation Council relations, are all examples of this style of international behavior. Although the results of this strategy remain meager, Europe can present itself as a new kind of actor, to distinguish it from the United States. In fact, the American government has a preference for the strategy of bilateral relations and the use of specific problem measures to manage international problems, while Europe has a strong preference for multilateral agreements and the global approach to manage cooperation programs with other countries. In the past 10 years, however, European governments have recognized that change in the area of international security, and also in the European region, demands the adoption of an assertive policy extending to military capabilities. The experience of the relations with Serbia on the occasion of the Kosovo crisis in 1995, after the negative experience of the European failure to manage the collapse of the Yugoslav federation, caused Europe to face security problems provoked by low-intensity conflicts that have had important destabilizing consequences in nearby areas. For this reason, Europe accelerated the formation of the common defense policy, launched the European security strategy, and established military and civil capabilities of crisis and conflict management as additional instruments of European international action.

This European resolve to act militarily in the form of crisis management and peace support operations is relevant also to the European presence in the UN. The European Community obtained observer status in the UN General Assembly in 1974. In subsequent years, this status was obtained also in many UN agencies, such as the Office of the High Commissioner for Human Rights (OHCHR), Office of the United Nations High Commissioner for Refugees (UNHCR), United Nations Conference on Trade and Development (UNCTAD), International Labour Organization (ILO), World Health Organization (WHO), United Nations Educational, Scientific and Cultural Organization (UNESCO), United Nations Environment Programme (UNEP), United Nations Development Programme (UNDP), United Nations World Food Programme (WFP), United Nations Office on Drugs and Crime (UNODC), and International Atomic Energy Agency (IAEA). Thanks to the exclusive competence the European treaties gave to the EU in various policy areas, the EU became a member of the Food and Agriculture Organization of the UN (FAO; as the European Community) in 1991, participates in many UN-specialized conferences, and has signed more than 50 conventions of the UN. On the whole, the importance of the EU and its member states in the UN system is remarkable. Although the 27 votes of the EU member states are one eighth of all the General Assembly votes, the contribution rate of the EU states to many areas of the UN system is larger than the vote rate. In the financial sector, for instance, EU states contribute about 38% of the UN ordinary budget, 50% of the contributions to special funds and programs, and 40% of the UN peace operations costs.

The EU's proactive attitude toward the UN showed a significant upward trend in the early 2000s. In 2001, the European Commission released the Communication on Building, an effective partnership with the UN in the fields of development and humanitarian affairs. In May 2003, a financial and administrative framework agreement was signed by the EU and UN. In 2003, the Commission released the communication on the EU and the UN: the Choice of Multilateralism, which made public the areas in which the Union was ready to engage itself as a UN partner. In September 2003, the European Council and the UN signed the Joint Declaration on cooperation in crisis management. This document is proof of the importance the EU governments give to ESDP as an instrument of European action in world affairs and multilateralism. Finally, it is worth remembering that in addition to the EU's role in economic and trade multilateralism, there is one other area in which European action in the UN and multilateral governance is very important. This is the area of environment protection, within which the EU has taken on a leadership role, even in confrontation with the United States, in many negotiations on crucial matters such as those relating to the ozone layer, climate change, desertification, and biodiversity.

In the UN agencies and programs, the EU presence has been the responsibility of the European Commission, but the role of the presidency has been very important in coordinating the position of the member states and in representing the EU in various forms, such as issuing official declarations and making démarches toward third countries. Actually, change in the responsibility of driving the European presence in the UN is an important example of the institutionalization of the EU transformation into an international actor. In fact, important reforms of this responsibility are contained in the Lisbon Treaty. The High Representative of the Union for Foreign Affairs and Security Policy will be given the function of coordinating the action of the member states in international organizations and conferences. The High Representative will have the right to be informed about matters of common interest by the EU member states, which are also members of the Security Council. In addition, when the EU has formed a common position on an object on the agenda of the UN Security Council, those member states that sit in the Security Council shall request that the High Representative be invited to present the EU's position.

The strategy of global approach of the EU as an international actor has not changed the importance of Europe in the world economic market and in other interactions. The EU is the area of the world with the largest internal market in terms of gross product and the largest actor in world trade. The European import–export rate is higher than that of the United States. Export contributes to the gross domestic product of the member states to an extent larger than that of the United States. These achievements, which are the result of the European economic integration process, cannot be safeguarded without appropriate actions in the world economic system. Therefore, in harmony with the liberal character of the European economy, the European external economic policy is aimed at preserving the solidity of the European economy within the free international trade regime. But preserving the strength of the European economy depends also on protectionist measures such as the external common tariff, the imposition of import quotas and antidumping measures, and the adoption of voluntary restrictions on exports. Accordingly, in the World Trade Organization, the EU supports the liberalization of world trade and the abolition of tariff restrictions to commerce. However, at the same time, it compromises on and delays the implementation of measures of trade liberalization when these measures may put European industries in difficult positions. Political and economic opportunity reasons also determine the definition of economic and trade agreements with single countries and groups of countries. Indeed, some economic agreements signed by the EU distort the rules of free trade for the sake of defending the European industries. This is the case, for instance, with the free exchange agreements with Israel and Switzerland; the mutual recognition agreements with the United States and Canada; and the trade agreements signed by the EU and countries such as Australia, Argentina, Brazil, Pakistan, and South Korea. Other trade agreements are aimed at developing good relations with nearby states—for instance, the European Economic Area agreement with Norway, Iceland, and Liechtenstein and the custom union with Turkey. Other commercial agreements aim at giving to the EU a politically rather than an economically favorable position in the dialogue with disadvantaged countries—for example, the Lomè Convention that grants preferential trade measures to 70 countries of Africa, the Caribbean, and the Pacific and the Generalized System of Preferences that gives to products coming from 145 developing countries free access to the European market. Other agreements aim at building good political relations with neighboring countries and giving preferential measures to economically weak countries, such as the Euro-Mediterranean Agreements with Algeria, Egypt, Jordan, Lebanon, Morocco, Palestine, Syria, and Tunisia, and the Partnership and Cooperation Agreements with 10 former Soviet Union countries. Finally, some agreements respond to the logic of regionalism of the contemporary world economy—for instance, the agreements with the Andean Pact, Southern Common Market (Mercado Común del Sur; MERCOSUR), the Central America custom union, the Gulf Cooperation Council, and the Association of Southeast Asian Nations (ASEAN). Briefly, the external economic action of the EU is driven by a policy that combines the strategy for consolidating the European position in the world trade system with strategies of a political nature aimed at either preserving existing political relations or creating new political solidarity with countries of interest to Europe. Both strategies interact with the sphere of the international political relations between the EU and external actors and are dependent on the ability of the EU to act as an international actor.

Conclusion

The EU is an international actor by ascription and practice. It has emerged to such a role thanks to interdependence and globalization and the self-change process of economic integration. In the 1980s and 1990s, European transformations in the form of the Helsinki Process and, later, the fall of the Communist bloc and the unification of regional politics promoted the search for a common security policy. Finally, the current reconfiguration of world politics has created the conditions for a self-conscious ascription of specific roles in the EU-constructed image of the world system. At the same time, it is true that the EU is something like an infant international actor that has to drive itself by trial and error in world affairs while facing the problems of building its own mechanisms of foreign policy making.

Fulvio Attinà
University of Catania
Catania, Italy

See also Diplomacy; European Integration; Europeanization of Policy; Foreign Aid and Development; Foreign Policy Analysis; Globalization; Governance, Global; Interdependence; Regional Integration (Supranational); Sovereignty; Transatlantic Relations

Further Readings

Bretherton, C., & Vogler, J. (2006). *The European Union as a global actor.* London: Routledge.

Ginsberg, R. H. (1989). *Foreign policy actions of the European Community: The politics of scale.* Boulder, CO: Lynne Rienner.

Laatikainen, K. V., & Smith, K. E. (Eds.). (2006). *The European Union at the United Nations: Intersecting multilateralisms.* London: Palgrave.

Smith, M. E. (2004). *Europe's foreign and security policy: The institutionalization of cooperation.* Cambridge, UK: Cambridge University Press.

Whitman, R. G. (1998). *From civilian power to superpower? The international identity of the European Union.* London: Macmillan.

European Integration

European integration is the voluntary creation and maintenance of regional institutions by states in Europe. It is legally binding and involves intense and varied patterns of interaction in areas such as economics, agriculture, justice, immigration, and foreign affairs. The creation of the European Coal and Steel Community (ECSC) in 1951 and the subsequent establishment of both the European Atomic Energy Community (EURATOM) and the European Economic Community (EEC) in 1957 marked the initial phase of European integration. Today, the European Union (EU) is the key pillar of European integration: It has 27 member states, and its constitutional basis lies in the 2009 Lisbon Treaty. EU institutions have binding powers in several fields, chiefly economic, and are burdened with corresponding levels of public expectation. However, the transfer of financial resources and political loyalty to the EU headquarters in Brussels, Belgium, has been limited. While the EU represents a new political community, largely autonomous from its individual member states, it has not superimposed itself on, or replaced, these states.

This entry discusses the entire process of integration, first, by giving the historical background at the origin of the very decision of creating the ECSC and the related complexities of the enterprise; second, by examining the key features of widening and deepening the process; third, by singling out the aspect of *EU* policy making; and fourth, by analyzing the role of the EU in world politics. A few remarks on the future perspective of EU conclude the entry.

Introduction

European history over the past 2 centuries was dominated by the forces of nationalism, left–right ideological battles, and the militarism accompanying these clashes of belief. The principal vehicle of modern European history up until 1945 was the nation-state: a territorially delimited entity whose predominant collective self-understanding was the nation. In his seminal history of Europe, A. J. P. Taylor treated European states as personified entities with human characteristics: French arrogance, British aloofness, and German frustration. One of the driving forces behind European integration since the early 1950s has been transformation of the nation-state in Europe. Looking at how states of the EU relate to each other today, we see that nation-states have been replaced by member states. These member states are committed to cooperation within the institutions and frameworks of European integration, a commitment that raises questions about their own sovereignty. Certainly, national self-interest has not disappeared, and states still bargain hard with each other over matters of policy. However, the 27 member states are committed to finding an agreement often at considerable cost to themselves. They also delegate to the EU powers that were traditionally thought of as the prerogative of sovereign states.

The very purpose and character of this exercise has never been clearly spelled out, and the end goals of the integration process have remained shrouded in mystery. From the very start of the European integration project, ambiguity, obfuscation, and the prevalence of competing sets of underlying principles has been the norm. Successive generations of European leaders have been faced with multiple and ever-changing agendas for which there were no readymade solutions. Persistent differences among European actors concern the very nature of integration (federalism vs. intergovernmentalism) and the functional scope of integration (high politics vs. low politics), and at each stage a compromise between competing national agendas has had to be found. In this situation, pragmatism, incrementalism, and vagueness become operative principles, with the result that ambitious and straightforwardly cooperative projects have had a fairly good chance of being shot down. A good example is the European Defense Community (EDC) launched by the French government in the early 1950s. Opposed by both Gaullists and Communists in the name of a dangerous federalism that would see French soldiers come under the authority of German officers, the EDC was voted down in the French National Assembly in 1954. Unable or unwilling to overcome these residual currents of national sentiment, European policymakers were faced with a choice: integration in disguise or no integration. The language of successive cooperative arrangements had to be vague, and no specific destination point for the European project was ever officially proclaimed. Although some scholars and politicians talked about constructing a European federation resembling the United States of Europe, most were happy to see the EU as an "unidentified political object," to use Jacques Delors's famous expression. Major decisions taken in the EEC and later the EU were thus based on a search for compromise and consensus rather than on any clearly articulated political principles or strategic end goals. The implications of this for democracy at the national level were never properly defined.

Deepening and Widening

In addition to such ambiguity pertaining to the ends of integration, the process itself has been an uneven mixture of deepening and widening. Extending cooperation to ever-new fields required different legal and procedural arrangements given the unwillingness of member states to delegate powers to the European Commission in certain areas. The same rules governing European competition policy or common agriculture policy could hardly be applied to the European justice and home affairs, let alone to the European Common Foreign and Defense Policy. Territorial widening of the integration project has generated yet further complexity. The EU has enlarged several times since the United Kingdom (UK), Denmark, and Ireland joined the original six member states of the EEC in 1973. With each new enlargement, there have been institutional and procedural changes in the EU system. New members have imported into the EU their diverse cultures, political concerns, and legal procedures. This has especially been the case with the "big bang" enlargement in 2004 that saw the entry of eight new member states from Central and Eastern Europe (followed by the entry of Romania and Bulgaria in 2007). These new member states are significantly poorer than their Western counterparts and are unlikely to close the wealth gap in the following decades. Their democratic institutions are still relatively unstable, and their economic, legal, and administrative structures are undeveloped when compared with those in the older member states. Central and Eastern European members also have their own distinct histories, societies, and cultures. For instance, while "de-communization" is a central preoccupation of political life in Eastern Europe, it plays no role in Western European politics. Eastern Europe still has relatively few immigrants from Third World countries compared with Western Europe, but it is struggling to come to terms with its own diverse and often sizable national minorities. These include Russians in Latvia and Estonia and Hungarians in Romania and Slovakia. New members have also different foreign and security priorities, especially regarding relations with the United States and Russia.

The EU's present complexity, however, is not the result of the enlargement process alone. The EU's diversity goes well beyond the simple East–West divide. It also corresponds to the different levels of development and of national experience across Europe as a whole. Countries such as Portugal, Greece, Spain, Ireland, and even Finland have all undergone significant modernization and adaptation after joining the European Communities/EU. The difficulties that arose in 2010 in Greece, Ireland, and Spain suggest that significant levels of unevenness pertain within the EU. In addition, Western European countries are struggling to overcome their own internal problems related to the incomplete integration of their large immigrant populations. Given the nature of the political responses generated by these challenges—from the emergence of populist rightwing parties to the spread of Euroscepticism across the continent—it comes as no surprise that the enlargement process has stalled. Though membership for Bosnia, Serbia, or Turkey might make good geostrategic sense, given how effective enlargement has been as a tool for stabilizing the EU's external environment, the entry of these countries remains in doubt. Burdened by the EU's own institutional, political, and social complexity, the future of its enlargement process is far from guaranteed.

Challenges to EU Policy Making

Such complexities indicate how far we have come since the days of the traditional mode of European governance, the so-called community method. This method rested on the presumption of equal rights and duties of all member states and the central role of the European Commission as guardian and agent of the common European interest. This method has repeatedly clashed with the forces of modernization and globalization, which generate increasing levels of interdependence in the fields of economic exchange and migration and the concomitant demand for discretionary and flexible responses by powerful national actors within the EU. That different fields of cooperation had evolved along with different memberships and different internal rules has made adaptation to changes all the more difficult. For instance, Norway and Iceland are parties to the Schengen Agreement, while the UK, Ireland, and Bulgaria are not, albeit for different reasons. Only some EU member states are members of the Euro Zone, but their decisions directly affect the entire EU. More flexible, diversified, and decentralized modes of governance have therefore been suggested for the EU, but it is far from clear whether they will prevail. This is partly because the experiments with these nonhierarchical types of governance have not been very successful. (The so-called open method of coordination, which relies on benchmarking of national progress, organized mutual learning, and deliberative problem solving across the EU, has failed to produce any tangible results.) This is also because the economic crisis that started in 2008 has reinforced the need for regulation and enforcement of the agreed rules.

As these challenges have grown, and alongside them the need to make crucial decisions pertaining to people's well-being, the dissociation between politics and policy making that was the hallmark of the Community method has become a major problem. The peculiar relationship between politics and policy making is at the heart of the "comitology" system. This system consists of some 250-plus committees set up to regulate the implementation of European Commission policies. The committees are chaired by the Commission, are composed of national officials from the member states, and are the site for detailed and highly technical deliberations over the secondary rules that govern how policies are to be implemented. In all political systems, negotiations over the implementation of policy agendas exist as a fundamental part of political life. What is distinctive about comitology is what Vivien Schmidt calls "policy without politics." Because the work of these committees is highly technical, they are staffed mainly by experts. The committees thus see themselves not as fora for political debates but as problem-solving bodies. Even when political questions intrude into the workings of these committees, this issue is mainly about the balance between the power of the Commission and the power of member states. Political battles amount to institutional turf wars rather than debates over political principles or ethics. Notwithstanding its central role in the organization of contemporary European society, the EU's mode of functioning thus excises the question of the "good life" from its deliberations. As a result, comitology is a system of policy making where the political questions about the EU's purpose and end goals are not addressed. This system has found itself under increasing pressure as the evidently political questions dealt with by the EU imply answers that have important consequences for people's daily lives.

At issue here is the question of democracy and public participation within the European integration project. Striking rejections of the EU's Constitutional Treaty in France and in the Netherlands in 2005 have been accompanied by a rise in Eurosceptic sentiment. Gone are the days of the "permissive consensus," where citizens of EU member states passively accepted the inevitability of an "ever-closer Union." Central to the debate about the "democratic deficit" of the EU is the question of standards. Against which standards should the EU be judged? Critics of the EU often imply that we should expect the EU to be as democratic as our own national political systems. Supporters of the EU respond that the EU is not a state and as such should be assessed on a different basis. As long as those decision-making powers delegated to the European Commission are adequately monitored (jointly by the Council of Ministers and the European Parliament), there is little cause for concern. Notwithstanding this debate about standards, certain facts remain. The European elections to the European Parliament are striking for their demonstration of the growing disinterest of European citizens. In the first direct elections to the European Parliament in 1979, turnout across the member states was 61.99%. Since then, there has been a steady decline: 58.98% in 1984, 58.41% in 1989, 56.67% in 1994, 49.51% in 1999, 45.47% in 2004, and 43% in 2009. Alongside the disinterest, doubts have been raised about the democratic quality of pan-European initiatives. In Germany, there was great reluctance to support Greece when the latter ran into economic difficulties in 2010. If national sentiment no longer corresponds to the processes and procedures of European integration, then realignment is necessary in order that the EU remain consistent with European public opinion.

From Peace in Europe to Peace Building Around the World

Throughout its history, European integration has been seen as an engine behind Europe's economic prosperity. Lifting exchange controls, eliminating quotas, reducing tariffs, and combating monopolist practices associated with the creation of the European Common Market have been generating economic competition and growth for some decades. However, as Giandomenico Majone and others have pointed out, growth has stagnated, or even regressed, since the launch of the two most important economic projects of the past 2 decades: the Single Market Programme and the European Monetary Union. European integration has not necessarily been responsible for this negative development. However, the lack of economic success has prompted European leaders to search for another rationale of European integration, and many have alighted on the role the EU can play in world politics. From tackling climate change to crisis management in conflict zones around the world, the EU is identified as a potential heavyweight. As China and India take their place on the world stage, it has become commonplace to repeat Jacques Delors's phrase, pronounced at a time when France had to decide on its European or national vocation: The choice is only between a United Europe and decline.

The EU is indeed, by all measures, a very powerful global actor. With its 27 member states, nearly 500 million inhabitants, a quarter of the world's GNP, around 40% of the world merchandise exports, and a comprehensive array of economic, legal, diplomatic, and military instruments at its disposal, the EU is able to exercise significant influence in various parts of the world. The euro is now the world's second most important international reserve and trade currency and as such serves as a conduit for EU influence across the global arena. Owing to the size of its internal market, European norms and regulations are progressively being adopted by others in the global economy. Leverage in this field allows the EU to shape the multilateral agenda on issues such as trade, investment, competition, banking, accounting, government procurement, and other types of services. Together, the EU and the United States produce around 80% of the international norms and standards that regulate global markets. The EU also campaigns for universal labor standards, environmental protection, eradication of poverty, and sustainable development. It champions the cause of international law and multilateral cooperation by supporting international endeavors such as the Convention on the Rights of the Child or the International Criminal Court. Democracy and human rights are also high on its agenda, even though the EU has eschewed the United States' more interventionist approach to democracy promotion and regime change. The EU, if one adds up national and EU-level funds, is also the largest provider of developmental aid.

There are other good reasons to take the European foreign policy project seriously. Diplomats from EU countries meet about 100 times a year and adopt more than 100 joint statements, communiqués, and declarations. Following the adoption of the Lisbon Treaty, the EU is now in the process of developing its own diplomatic service. The EU's institutional framework has become the focal point for European foreign policy debates, and it is the preferred point of departure for most national diplomacies. Today, all EU member states try to speak and act "in the name of Europe," even if that does not mean acting through Europe itself. In recent years, the EU's contribution to international peace and security has seen it dispatch missions to far-flung places, from East Timor, Congo, Sudan, and Afghanistan to Iraq, Lebanon, Bosnia, and Georgia.

However, the EU's global policy is confronted with numerous problems. Europe's external trade relations are largely divorced from Europe's foreign policy. Responsibility over external trade is shared or split between the European Commission, the European Central Bank, the Council of Ministers, the Eurogroup, or the member states. Foreign policy proper, unlike external trade, is subject to individual countries' veto. Moreover, EU member states are allowed to act outside the EU framework, and, in fact, they frequently do so, either within the United Nations framework or via the Organization for Security and Co-operation in Europe, the Council of Europe, or NATO. European foreign and security policies are often carried out by formal or informal coalitions of the willing, by contact groups or bilateral initiatives. And the EU has practically no military standing force at its disposal and has to rely on member states' voluntary contributions.

Export of EU laws and norms has proved successful chiefly in the EU's immediate neighborhood where its power was overwhelming and its norms were shared, as was the case in Central and Eastern Europe. In regions farther away, the EU's efforts to promote its norms in fields such as trade, environment, or social protection are often seen as disguised forms of protectionism, of benefit mainly to those European producers and consumers wishing to be shielded from the competitive pressures of less developed countries. In fact, the opposition of a group of developing countries led by India and Brazil has made it impossible for the EU to include as part of the multilateral trade agenda the so-called Singapore issues: competition policy, investment, government procurement, and trade facilities. Moreover, major players such as the United States or China are able to play individual EU member states against each other, effectively undermining the EU's aggregate potential. This is because the EU lacks a European equivalent of the *raison d'état*: As the president of the European Council, Herman Van Rompuy, put it, the EU's strategic partnerships with rising powers such as China lack any real strategic vision. Moreover, the EU's mechanisms for identifying common interests are weak and unable to overcome national particularities, leaving it susceptible to divide-and-rule practices by rivals such as Russia.

Conclusion: Twilight of the EU?

European integration has never been a smooth or linear process, and it has always had its critics. Nevertheless, in recent years doubts have been raised about the very viability of the EU. Its political legitimacy was profoundly shaken by the failure of the European Constitution in 2005 and the arm-twisting that occurred at the time of Ireland's ratification of the Lisbon Treaty in 2008. Its existence as a stable and unified monetary community is in doubt as weaker member states—from Ireland to Portugal and Greece—struggle to service their debts without external help. The EU's foreign and security policy is a continual source of frustration for those who believe in the EU's vocation as a power with global reach; from the Balkans to the Middle East, the EU is still unable to credibly answer Henry Kissinger's often-cited question about which number to call when wanting to get through to "Europe."

These crises and disappointments have clearly undermined the EU's credibility and stalled individual projects either in the domain of widening or of deepening. And yet as in previous moments of instability and crisis, it would be premature to talk of the swing of the pendulum from integration to disintegration. As politicians of EU member state governments themselves admit, European integration is the context within which they all think and act. Calls to exit the EU are still viewed as quixotic; it is rather like wanting to retreat from a globalized and interconnected world in favor of a premodern idyll. The EU's central function as a mediator of conflict between its members suggests that in the years to come it may be more necessary than ever. Moreover, the structural shift from nation-states to member states implies that Europe itself has changed and that European integration has become an integral feature of contemporary statehood today in Europe. The end of European integration also runs against the play of international politics as it would make it difficult for European states to aggregate their power when faced with global competitors such as China, India, Russia, or the United States of America. There may be much to criticize about European integration, but alternatives to it seem either utopian or outdated. Adjustment to the rapidly changing external environment and the management of internal public expectations was and still is the greatest challenge for the EU in the years to come.

Chris J. Bickerton
University of Amsterdam
Amsterdam, Netherlands

Jan Zielonka
University of Oxford
Oxford, United Kingdom

See also Autonomy, Subnational; Comparative Politics; Cross-National Surveys; Europeanization of Policy; Governance, Multilevel; Intergovernmentalism; Parties; Regionalism; Welfare State

Further Readings

Bickerton, C. J. (2011). *European Union foreign policy: From effectiveness to functionality.* Basingstoke, UK: Palgrave.

Hay, C., & Menon, A. (2007). *European politics.* Oxford, UK: Oxford University Press.

Hill, C., & Smith, M. (2005). *International relations and the European Union.* Oxford, UK: Oxford University Press.

Hix, S. (2010). *What's wrong with the European Union and how to fix it?* Cambridge, UK: Polity.

Leconte, C. (2010). *Understanding Euroscepticism.* Basingstoke, UK: Palgrave.

Majone, G. (2009). *Europe as the would-be world power: The EU at fifty.* Cambridge, UK: Cambridge University Press.

Milward, A. (2000). *The European rescue of the nation state.* London: Routledge.

Moravcsik, A. (1999). *The choice for Europe: Social purpose and state power from Messina to Maastricht.* London: UCL Press.

Rosamond, B. (2000). *Theories of European integration.* Basingstoke, UK: Palgrave.

Schmidt, V. (2006). *Democracy in Europe: The EU and national polities.* Oxford, UK: Oxford University Press.

Wallace, H., Wallace, W., & Pollack, M. (Eds.). (2005). *Policy-making in the European Union.* Oxford, UK: Oxford University Press.

Zielonka, J. (2006). *Europe as empire: The nature of the enlarged European Union.* Oxford, UK: Oxford University Press.

Europeanization of Policy

Europeanization most often refers to domestic adaptation to regional integration. It can also denote the diffusion of European values and forms of governance to other parts of the world. In a third meaning, the concept may refer to the emergence of common political standards and patterns throughout Europe. Studies on Europeanization focus on different aspects of the political system: values, institutions, policies, policy styles, and policy procedures. There is no consensus among scholars as to what, more specifically, Europeanization is and where it takes place. This entry focuses on two different ways of defining Europeanization as adaptation to regional integration, because these are the most frequently used definitions. One popular definition is that Europeanization is the impact of the European Union (EU) on domestic policy, polity, and politics. This implies that Europeanization is something that happens within EU member states and others, and Europeanization is thus seen as an outcome of European integration. Another definition derives Europeanization from the interaction between European and domestic policy processes. In this view, Europeanization is seen as a process embedded in the evolution of European integration. These two different ways of defining Europeanization are accompanied by different sets of analytical concepts and generate different research questions.

Europeanization is not a theory of its own, but the increasing number of studies of Europeanization since the end of the 1990s signals a growing interest in the theoretical aspects of the integration process. Even though the intellectual development of this concept has not been particularly tidy, it shows a shift in focus by students interested in theories of European integration from the development of laws, regulations, and institutions on the EU level to changes at domestic levels. In this way, the concept has been used not solely for studies of EU member states but also for states that are not EU members (both future members and other states).

Europeanization as an Outcome

When Europeanization is seen as the impact of the EU on domestic policy, polity, and politics, it conceptualizes Europeanization as taking place on the domestic level. In this model, Europeanization is defined as national or regional changes caused by European integration. Through this prism, one can study Europeanization as the consequence of a formal or informal European integration process. In this model, the main research question is frequently whether, or to what extent, something has been Europeanized.

In this understanding of Europeanization—as used, for example, in an anthology edited by Maria Green Cowles, James Caporaso, and Thomas Risse and in a much-cited article by Robert Ladrech—one needs to specify which specific domestic changes in policy, politics, and polity can be caused by European integration and can count as Europeanization. The literature on the concept does not agree on the relevant effects of the process. Analyses of transposition and implementation examine changes in legislation and legal practice, whereas other studies focus on policy change, institutional persistence, and policy discourse.

The literature also examines the mechanisms by which this impact comes about, and important efforts have been made, for example, by Claudio Radaelli, Christoph Knill, and Dirk Lehmkuhl to systematize how domestic changes take place, drawing on the conceptual toolboxes of various disciplines. In some cases, there are sequences of voluntary Europeanization in one area as a spillover of coercive Europeanization in another policy area.

Another distinction has been made between vertical and horizontal mechanisms of Europeanization. Vertical mechanisms function where there is a European model that a state should adhere to, resulting in direct adaptational pressure. The horizontal mechanisms involve different kinds of framing processes, such as learning, inspiration, mimicry, and copying. From this point of departure, the mode of governance in a policy area may help predict the character and efficiency of the Europeanization drivers. This means that the organizational placement of a policy area as either a supranational policy area (where law is made) or an intergovernmental policy area (where policy primarily is coordinated) may influence the levers of Europeanization.

Europeanization as a Process

Even in policy areas where there is no supranational policy process or where its impact is very limited, domestic policy is influenced by what happens in other European states and more generally in the European context. In this way, the interaction between European and domestic policy processes can be thought of as generating new policy areas.

The novelty in these policy areas lies less in substance than in shape and framing. When previously domestic policy areas are affected by European policy making and jurisprudence, they often change names and demarcation lines to other policy areas. Such shifts occur not only in fields where the EU has legislative competence but also in many other spheres. It might be more accurate to speak of partially Europeanized policy areas. Both state and nonstate actors within such areas have a growing number of international contacts and are increasingly exposed to European rules and other extraneous influences. The process of adaptation is seldom unidirectional and sequential. More often, it takes the form of gradual and mutual adaptation, with increased integration and domestic changes affecting each other and taking place in parallel.

While some studies treat Europeanization as an objective phenomenon, others emphasize its rhetorical and discursive dimensions by treating it as a construct with particular political functions. A popular research design makes Europeanization a dual process in which European and domestic political patterns are intertwined. Some scholars, for example, Ulrika Mörth, argue that it is impossible to study the European and domestic processes in separation as they are increasingly interwoven and dependent on each other. Others, for example, Ian Bache, Simon Bulmer, and Martin Burch, have examined Europeanization as a two-way process and made efforts to distinguish the different streams from each other. From this point of view, Europeanization not only consists of the downloading of EU policies that affect member states' policy, politics, and polity but also of its uploading.

Conceptualizing Europeanization as an interaction between domestic and European levels makes it difficult to measure the effects of the integration process. Recent researchers tend to distance themselves from the sequential conception of policy making in which decision makers were seen as identifying problems, setting goals, proceeding then to compare alternative strategies, and choosing those with a favorable ratio between costs and expected results. Instead, following John Kingdon (building on the garbage can model), they operate at the intersection of problems, policies, and solutions floating around in the partially Europeanized policy processes.

The question of when Europeanization takes place is also complicated by the temporal aspects of the European integration process that in several ways affect domestic planning horizons. Within this model, the important research question is not whether Europeanization has taken place but how it proceeds and manifests itself in various modes of interaction between European and domestic policy processes.

Of crucial importance for this line of research is a closer look at the specific linkages and ligatures of multilevel governance. Understanding the nature of Europeanization requires notions illuminating the particular ways and forms in which European and domestic policy processes are related. Some concepts in the literature on Europeanization that might guide such an analysis are customizing, framing, the idea of social representations, and mediated cross-load.

Customizing refers to efforts undertaken by national politicians to make the EU resemble their own state. When participating in the EU policy process, member states emphasize features of the European policy that makes it more similar to the domestic policy, thereby making the EU more familiar to various national constituencies. This is a way of strengthening the legitimacy of the EU within the national political culture.

A related notion is that of framing, which stands for the context in which a policy is presented. Again, the purpose is to gain acceptance for specific proposals or lines of action. A similar idea is that of "social representations," which shape the institutionalization of the policy area and define frameworks for possible cooperation.

An idea related to that of upload and download discussed earlier is that of cross-load, where local practices are disseminated from one state to another without necessarily passing the central EU level and its institutions in Brussels. However, even if practice is not disseminated via the EU institutions, they are not necessarily isolated policy transfers, but they can be expected to be mediated by other processes on the European level, which would be mediated cross-load. Such processes might take place in transnational networks. In addition, policy transfer in a policy area may take place between some countries and might then be pushed "upward" to the European policy process. Such conceptual tools have been employed in the analyses of Europeanization encompassing domestic changes that precede and facilitate further formal integration.

Daniel Tarschys
University of Stockholm
Stockholm, Sweden

Malena Britz
Swedish National Defence College
Stockholm, Sweden

See also European Integration; Globalization; Implementation; Institutionalization

Further Readings

Bache, I. (2008). *Europeanization and multilevel governance: Cohesion policy in the European Union and Britain.* Lanham, MD: Rowman & Littlefield.

Bulmer, S., & Burch, M. (2001). The Europeanisation of central government: The UK and Germany in historical institutionalist perspective. In G. Schneider & M. Aspinwall (Eds.), *The rules of integration: Institutionalist approaches to the study of Europe* (pp. 73–96). Manchester, UK: Manchester University Press.

Featherstone, K., & Radaelli, C. (Eds.). (2003). *The politics of Europeanization.* Oxford, UK: Oxford University Press.

Graziano, P., & Vink, M. P. (Eds.). (2007). *Europeanization: New research agendas.* New York: Palgrave Macmillan.

Green Cowles, M., Risse, T., & Caproaso, J. (Eds.). (2001). *Transforming Europe: Europeanization and domestic change.* Ithaca, NY: Cornell University Press.

Ladrech, R. (1994). Europeanization of domestic politics and institutions. *Journal of Common Market Studies, 32*(1), 69–88.

Evaluation Research

Evaluation research is about assessing the consequences of public policy. The emphasis can be either on the policy process or on the outputs and the outcomes of the policies. Evaluation research thus has a clear democratic and learning goal: to inform the decision makers, citizens, and interest groups whether public policies have or have not had the intended consequences. Whether public policies result in alleviating public problems is a central question for many societal stakeholders. Governments are keen on showing that the practiced policy making has been effective and that promises made during an election campaign have been kept. In addition, evaluation is an important tool of monitoring; governments as principals use evaluation to keep the agents and the bureaucracy accountable. Opposition parties, in principle, are as keen on showing that government policies have not had the intended effects. The media and think tanks inform citizens, and to do so, they need information on activities. This definition implies, first, that evaluation research is usually empirical and, second, that evaluation is usually bound to the goals of public policy making. This is, however, not the only option. Seen against the multitude of benefits it serves, evaluation research is not without problems. The academic element of evaluation research is to develop and further improve the tools with which to evaluate. The problems of evaluation concern both the practice of evaluation—how to do it—and the usage of evaluation. This entry focuses on the how question, and thereafter the focus moves on to the usage of evaluation.

Conducting Evaluation

Researchers engaged in the evaluation of public policies face a number of choices related to approach, focus, and criteria. First, various forms of evaluation research differ from each other on the basis of the selected approach. The principal approaches are formative and summative evaluation. Formative evaluation monitors the way in which a program is being administered or managed so as to provide feedback that may serve to improve the implementation process. Summative evaluation measures how a program has actually affected the problems it was designed to address. Second, evaluation can focus on various parts in a policy process: on the process or implementation at various levels of government, ranging from the top of the hierarchy to the grassroots bureaucrats, or on the outputs and outcomes. Third, evaluation can focus on various aspects of the process or outputs. The approach and focus used determine the criteria for evaluating a program's success. Usually, evaluation involves a comparison of the goals and impacts of a program, with the program's goals often selected as the yardstick for measuring or evaluating success. But efficiency, organizational changes, stakeholder concerns, and democracy can also be used as criteria. In other words, if a new policy aims at diminishing traffic accidents, the evaluator assesses whether this indeed has been the result. Other options include focusing on economy (Has money been wasted?), democracy (Did the program respond to popular will?), or impacts. In recent decades, public administration has moved into using performance indicators (PI) to guide agencies and lower rank officials to aim at the negotiated targets. For example, in academia, PI may include the number of publications and student achievement scores; for health care, PI may include the number of operations and patient calls. The debate concerning the use of PI has pointed out that the choice of indicators is important so as not to narrow the target field too much and has warned that top-down set indicators can have a demoralizing effect on lower rank officials' professional drive for work improvement.

The formal goals of policies form an important starting point for evaluators. This is because government ought to respond to the needs and expectations of the people. If evaluation shows that the results are poor or that the way in which programs are carried out is not optimal, government should be concerned. At the same time, there are a number of reasons to defend the replacement of formal goals as an evaluation criterion. First, political goals are not always very precise. Public programs can contain multiple goals, some of which may be meant to satisfy the voters rather than target a problem. Second, concentrating solely on monitoring the realization of formal goals may restrict the identification of various side effects that policies may have. These are usually not evaluated as thoroughly as the formal goals. Yet unintentional side effects can replace and disturb the formal goals in the long run.

An extreme option for criteria is to select a goal-free approach—that is, to study a policy process from all possible angles. A more recent approach, a realistic one, asks what works for whom in what circumstances, emphasizing thus the importance of the context in which a policy is making an impact. All in all, the various evaluation models or approaches have various emphases and hence may lead to different sorts of conclusions as to what constitutes a successful program. An evaluation researcher usually either chooses a particular perspective, say how a regional policy affects the participation of the local citizens, or can aim at reconstructing a comprehensive account of the merits of a program. The latter choice inevitably involves a selection of multiple criteria, data, and perspectives in a program. Finally, as in social science in general, some researchers prefer a deductive approach and others an inductive one; some prefer a quantitative analysis, others a qualitative case approach.

Whereas evaluating the administrative process whereby policies are fulfilled can be relatively easy, emphasizing the impact of public policies is loaded with difficulties. This is mainly because public policies are not the only factor that affects societal processes. A standard example is education. Although schools play a vital role in education by providing the institutional framework (schools, teachers, and resources) for learning, education is also dependent on the pupils themselves, their parents, and the underlying culture. In other words, a number of other factors may intervene once a public policy is implemented. Regional policy offers a good example: A tax relief program may be proposed to encourage private sector enterprises to add jobs. Once the program is implemented, it is but a single factor when such enterprises consider whether or not to invest in hiring additional staff. In education, similarly, good education is a necessary but not a sufficient ingredient of learning. Full impacts take time to be realized. In other words, when public policies are put into effect, as time passes, it becomes difficult to distinguish between the net effect of the policy and other effects. Furthermore, the task is not made easier by the fact that the grassroots officials often cooperate with other officials. Environmental health inspectors, for example, talk and meet with representatives of other municipal branches (garbage, health, fire, and planning), other public officials and relevant private shopkeepers, producers, and so on, and all this adds up to environmental health. One particular policy constitutes only a fraction of the total impact; alone it does not suffice to make much difference. A bottom-up perspective into a policy process opens up a different scenario compared with a top-down perspective.

Use of Evaluation Results

The usage of evaluation results is also not without problems. First, the aforementioned methodological constraints mean that evaluation results are not totally reliable. One can, for example, argue for or against the application of one specific criterion or approach. The evaluation of economic efficiency, for example, can be criticized if it does not include in the analysis an indicator for the quality of services too. The same applies to other criteria, and hence, one solution is to apply several criteria. Also, it may simply be impossible to sample all the necessary data. Evaluation as a part of monitoring usually has to rely on quantitative data. European Union programs, for example, are assessed on the basis of the number of new jobs created. There is, however, no certainty that these jobs will be lasting. Qualitative data may, however, be more difficult and costly to obtain in a day-to-day routine evaluation. Second, the difficulties of impact assessment are serious but not impossible. It is evident that public policy outputs need to be balanced in relation to other intervening factors. In the recent Programme for International Student Assessment (PISA) evaluations, which compare international education results, reading, for example, is defined as determined by a combination of teaching, the pupil's own activity, and support by parents. Impacts are also complex in the modern world. Globalization, crowded policy space, and postmodern citizen attitudes mean that the way individual policies influence the target problems is more and more uncertain. However, full impacts can hardly be assessed on a routine basis. Also, as the realistic argument says, the contextual nature of policy implementation implies that the process and hence also the impact varies, and explanations do not rest on certainty but on varying degrees of plausibility.

Irrespective of how skillfully and comprehensively an evaluation is conducted, its results may be poorly used. This has to do with power. Governments and other policymakers may not be willing to publish news when the impact of government programs has been poor, nor will the responsible agency applaud if the results of its activities are criticized. After all, public agencies are interested in both good results and maintaining their status. Evaluation results are, in other words, treated strategically, and this can take various forms. Evaluation can, for example, be designed and conducted in concert with the individual or organization who commissions it. Thus, the selection of criteria can be affected by factors other than the principles of good evaluation. A second possibility is simply that evaluation results may be rejected if they do not correspond with the wishes of those who authorized the evaluation. Evaluation is a part of democratic governance. Feedback and openly discussing impartial evidence are signs of transparency and openness. A perfect evaluation is seldom possible due to the aforementioned restrictions. Evaluation should not be abandoned, and the effort to see how public programs are working should not be given up. Evaluation is just one element in the debate over policies and programs. It may have an influence but so do other factors, including political commitments, the preservation of institutional relationships, public opinion, and so on. Politics in democratic societies, on the other hand, is about the belief that policy making can and will make differences, and hence public policy making needs feedback. There is a lot of uncertainty involved. Successful evaluations are based on clear, explicated criteria. A particular debate within evaluation research concerns the underlying theoretical assumptions of the various approaches and to what extent these affect the evaluation results. Evaluation research, in other words, cannot escape the common basis of social science to which it belongs.

Pekka Kettunen
University of Jyväskylä
Jyväskylä, Finland

See also Accountability; Monitoring

Further Readings

Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.

Patton, M. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
Pawson, R., & Tilly, N. (1997). *Realistic evaluation.* London: Sage.
Rossi, P., & Freeman, H. (1999). *Evaluation: A systematic approach* (6th ed.). London: Sage.
Vedung, E. (2000). *Public policy and program evaluation.* New Brunswick, NJ: Transaction Publishers.

EVENT COUNTS

Many subjects of interest to political scientists take the form of event counts: nonnegative, integer counts of the number of times that a particular event has occurred. Event count data are found in American politics (e.g., counts of the number of policy adoptions, presidential vetoes, and judicial decision reversals), international relations (e.g., interstate conflicts, militarized disputes, and treaty violations), and comparative politics (e.g., protest activities, cabinet changes, and regime transitions). Just as scholars are often interested in estimating the time of event occurrence, via event history analysis, so also are they often interested in modeling counts of events that are observed, via event count models. Count data present problems for the linear regression model developed for continuous dependent variables, which can produce biased, inconsistent, and inefficient estimates when applied to count data. As a consequence, scholars have developed several alternative event count models for count data. These are discussed below.

Poisson Regression Model

The most basic event count model is the Poisson regression model, which assumes that the count data are produced by a Poisson process, with a conditional mean captured by the covariates in the model:

$$\Pr(y_i|x_i) = \frac{\exp(-\mu_i)\mu_i^{y_i}}{y_i!}, \qquad (1)$$

where $\mu_i = \exp(x_i\beta)$. Because the Poisson regression model is a loglinear model, the expected count changes by a factor of $\exp(\beta_k)$ for a one-unit change in x_k. The Poisson regression model can be estimated by maximum likelihood. Although it is natural to conceive of count processes as following a Poisson process, the Poisson regression model actually contains three quite restrictive limitations that limit its applicability. First, the Poisson process is premised on events being independent of each other—the occurrence of an event does not increase the probability of the unit experiencing the same type of event in the future. This assumption will often be violated in practice. Second, related to this, the Poisson regression model also assumes that the conditional variance is equal to the conditional mean, a condition known as equidispersion. In many applications, the conditional variance will exceed the conditional mean, a condition known as overdispersion. Third, the Poisson regression model underpredicts the number of zeros (absences of events) that are observed in many applications. More sophisticated count models than the Poisson regression model account for these divergences from the assumptions of the Poisson model.

Overdispersion

In an event count model, the dependent variable is the number of events, not the occurrence of an event. As a consequence, the researcher needs to make assumptions about the process that aggregates to the event count. When overdispersion is present, estimates from the Poisson regression model generally will still be consistent if the conditional mean is properly specified. (Note, however, that truncated or censored data can result in inconsistency, even if the conditional mean is properly specified). Standard errors from Poisson regressions, however, will be biased downward, leading to Type I errors (rejecting the null hypothesis when it is true) in inference. As a consequence, overdispersion must be incorporated into the model specification when present.

There are two principal sources of overdispersion in count data. On the one hand, unmodeled sources of heterogeneity may produce a conditional variance that is larger than the conditional mean. On the other hand, the assumption of independence of events inherent in the Poisson process may be violated. Typically, researchers conceive of this nonindependence as a form of contagion within units—the experience of an initial event increases the probability of the same unit experiencing subsequent events. For example, a country's experience of an interstate conflict may increase the probability of the country experiencing subsequent interstate conflicts.

Given the tendency of event count data to exhibit overdispersion, it is important to test for this violation of the Poisson regression assumption rather than simply accepting the estimates from a Poisson model. A variety of tests for overdispersion have been developed. A common regression-based test for overdispersion typically assumes that the overdispersion takes the following form:

$$V(y_i|x_i) = \mu_i + \alpha g(\mu_i), \tag{2}$$

where α is the dispersion parameter and g is a known function. Under the null of equidispersion, $\alpha = 0$. A test of H_0: $\alpha = 0$, then proceeds by estimating a Poisson regression model and running an auxiliary ordinary least squares (OLS) regression. If the null of equidispersion can be rejected, the researcher will then wish to estimate a model capable of capturing overdispersion.

Negative Binomial Regression Model

The most commonly employed model for overdispersion is the negative binomial regression model. The mean $\mu = \exp(x\beta)$ from the Poisson regression model is replaced in the negative binomial regression model with a random variable:

$$\begin{aligned}\tilde{\mu}_i &= \exp(x_i\boldsymbol{\beta} + \varepsilon_i)\\ &= \exp(x_i\boldsymbol{\beta})\exp(\varepsilon_i),\\ &= \mu_i \exp(\varepsilon_i),\end{aligned} \tag{3}$$

where ε_i is an error term reflecting unobserved heterogeneity. In other words, the heterogeneity in unmodeled covariates that produces a variance that exceeds the mean is modeled in the negative binomial regression through the random variable $\tilde{\mu}$. With $\delta_i = \exp(\varepsilon_i)$, the distribution of the y_i conditional on x_i and δ_i remains Poisson:

$$\Pr(y_i|x_i, \delta_i) = \frac{\exp(-\tilde{\mu}_i)\tilde{\mu}_i^{y_i}}{y_i!}, \tag{4}$$

Typically, a gamma distribution with parameter θ_i is assumed for the δ_i, resulting in the following negative binomial probability distribution:

$$\Pr(y_i|x_i) = \frac{\Gamma(y_i + \theta_i)}{y_i!\Gamma(\theta_i)}\left(\frac{\theta_i}{\theta_i + \mu_i}\right)^{\theta_i}\left(\frac{\mu_i}{\theta_i + \mu_i}\right)^{y_i}, \tag{5}$$

The expected value, μ_i, remains the same for the negative binomial distribution as in the Poisson. The conditional variance in the negative binomial, however, is larger than the conditional mean:

$$\mathrm{Var}(y_i|x_i) = \exp(x_i\beta)\left(1 + \frac{\exp(x_i\beta)}{\theta_i}\right). \tag{6}$$

The negative binomial regression model can be estimated by maximum likelihood.

Excess Zeros

A variance larger than the mean, overdispersion, is evidenced by a larger number of zero counts of events (the absence of any events) than is predicted by the Poisson regression model. The negative binomial regression model accounts for this excess number of zeros by modifying the conditional variance from the standard Poisson model. Heterogeneity and contagion, however, are not the only possible sources of excess zeros. An excessive number of zeros may also be produced by a dual-regime process, in which the data-generating processes producing zero counts and producing event counts once events are possible are qualitatively different. The first of the two stages is a transition stage, which governs the probability of transitioning from experiencing no event to experiencing events. The second stage is the events stage, which produces the number of events that are experienced once the unit has transitioned into the stage in which events are possible.

If a dual-regime process is suspected, rather than modifying the conditional variance and leaving the conditional mean unchanged from the Poisson model, the researcher will instead wish to modify the model for the conditional mean. This is done by recognizing the dual regime nature of the data-generating process. Two principal dual-regime alternatives have been proposed: the zero-inflated model and the hurdle model. The two alternative dual-regime models differ in their conception of the possibility of zero counts once a unit has exited the transition process.

Zero-Inflated Models

In zero-inflated models, zero counts can be observed in either of two distinct regimes. In the first regime, the outcome is always zero. In the second regime, events or nonevents may be observed. Although developed for the application of identifying defects in manufacturing processes (in a perfect manufacturing process, defects are never observed in the first regime), it is easy to think of political science applications in which these two distinct types of regimes are possible. For example, under certain conditions, states may never go to war; however, if these conditions are not met, then either war or the absence of war may be observed.

The probability of exiting the regime in which only zeros are observed can be parameterized as a logit or probit. Poisson and negative binomial models have been used to parameterize the second events stage. But whether one uses a Poisson or negative binomial specification, the commonality of these zero-inflated approaches is that zero counts are possible for either of the two regimes in the model. This latter feature sets these models apart from another common event count model, the hurdle model.

Hurdle Models

Like their zero-inflated counterparts, hurdle models also assume a dual regime process. Hurdle models, however, depart from zero-inflated models in assuming that zero counts are not possible in the events stage. Instead, once the unit has crossed the hurdle and departed the first stage, at least one event will be observed. The second-stage process is thus truncated at zero. As a consequence, the probability of excessive zeros is determined solely by the first-stage process. As with zero-inflated models, both Poisson and negative binomial versions of hurdle models are used by researchers, depending on their conception of the events stage process.

Conclusion

Political scientists are increasingly interested in modeling phenomena for which the standard linear regression model is not appropriate. One of these phenomena is count processes, which arise in each of the empirical subfields of the discipline. A variety of specialized models have been developed to model count processes, and research continues on these models. Important recent developments in event count models include correlated event count models, panel data models, and Bayesian models.

David Darmofal
University of South Carolina
Columbia, South Carolina, United States

See also Event History Analysis; Statistical Inference, Classical and Bayesian; Maximum Likelihood

Further Readings

Cameron, C., & Trivedi, P. (1998). *Regression analysis of count data*. Cambridge, UK: Cambridge University Press.

King, G. (1989). Event count models for international relations: Generalizations and applications. *International Studies Quarterly, 33*(2), 123–147.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables.* Thousand Oaks, CA: Sage.

Winkelmann, R. (2003). *Econometric analysis of count data* (4th ed.). Berlin, Germany: Springer-Verlag.

Zorn, C. J. W. (1998). An analytic and empirical examination of zero-inflated and hurdle Poisson specifications. *Sociological Methods & Research, 26,* 368–400.

EVENT HISTORY ANALYSIS

Event history analysis allows researchers to investigate not only occurrences of certain political events but also the processes or histories of those events. In essence, event history analysis estimates how histories of events or durations leading up to the events (if events occur at all) change with different values of independent variables (or covariates). Such estimates allow researchers to discuss how included covariates increase or decrease the probability of the event occurrence at any given time. Since event history analyzes the duration between the start when the event is possible to the time when the event actually occurs (if it occurs), event history analysis is often called *duration analysis*. Alternatively, as event history analysis models the probability of an observation surviving (not experiencing the event), event history analysis is sometimes called *survival analysis*. Below, this entry discusses some major features and possible applications of event history analysis, taking into account both parametric and nonparametric models and some recent extensions.

Many political science studies ask questions regarding occurrences of significant political events. For instance, scholars in comparative politics are interested in why some countries democratize while others fail to do so (democratization), or why some countries are able to economically develop while others stay in poverty (economic development). Likewise, scholars in international relations study why some countries join many international organizations, while others join only a few (international organization [IO] membership) and why some pairs of states are more likely to engage in militarized interstate disputes than others (militarized disputes). Scholars in American politics ask why some court nominees are approved, while others are not (legislative and judicial branch relations), and what makes an incumbent more likely to face a challenger (reelection). By answering these questions, political scientists are able to identify systematic dynamics influencing occurrences of these political events.

While these inquiries certainly enhance our understanding of interesting political phenomena, the timing of occurrences of such events can provide even richer information and thus help further enhance our understanding of the underlying dynamics. For instance, not only identifying under what conditions a country is able to democratize but also finding out under what conditions a country is able to do so faster than other countries would help promote our understanding of democratization. Event history analysis provides researchers with a unique opportunity to exploit richer information of the histories of events.

Since event history analysis has been introduced to political science, there have been a growing number of studies that use it in empirical investigations. With this powerful new tool, many old research questions have been revisited, and many new research questions have been developed and empirically tested. These questions cut across subdisciplines of political science, and examples of them are numerous. Why do some coalition governments last longer than others? Why do some militarized disputes last longer than others? Why do some periods of peace last longer than others? Why do some economic sanctions last longer than others? Why do some drugs take a longer time to get approved by the Food and Drug Administration than other drugs? And why can some representatives maintain their seats longer than others? The popularity of the event history analysis is due to its usefulness in understanding the underlying political dynamics of events and their histories. Specifically, event history analysis has a few advantages over traditional binary logit or probit analysis, on the one hand, and over ordinary least squares regression (OLS), on the other.

Event history analysis has many advantages over binary logit or probit models. First of all, event history analysis takes the history of an event seriously by taking duration into account. As how long it takes for an event to occur provides more information than just whether the event occurs or not, it helps better understand the dynamics of the events under investigation. Returning to a prior example, some countries may successfully democratize less than a few decades after independence, others may democratize after more than a few decades, while still others may not be able to democratize at all. Conventional logit or probit models cannot differentiate between those cases democratizing in less than a few decades and those in more than a few decades. In contrast, event history analysis allows researchers not only to investigate factors that make democratization more likely but also to investigate factors that hasten democratization by exploiting richer data.

Likewise, one may think that OLS regression might be able to capture factors influencing duration quite nicely, as it examines continuous dependent variables. But event history/duration data pose several challenges for traditional OLS estimation. For one, duration data are often right skewed and the OLS approach requires an arbitrary transformation of data (most commonly taking a log) to deal with a right-skewed data set. A more serious problem is data truncation. Data truncation happens when researchers do not know either the exact entry time of an observation (left truncated) or the end time—this is the case when events have not yet occurred when the data collection is done (right censoring). That is, there could be some observations that, at the time of data collection, are still ongoing. OLS treats left-truncated observations as if they have an equivalent entry time with other observations and copes with right-censoring problems either by dropping all the observations that have not experienced the event or by capping the history by assuming that the event has occurred at the conclusion of the period of data collection. Both these arbitrary assumptions can cause biases. In contrast, event history analysis takes into account these cases of left truncation and right censoring. Finally, event history analysis can accommodate time-varying covariates (TVCs), which are independent variables with different values over time.

In event history, an "event" is the primary phenomenon of interest, and the "history of such an event" refers to the duration leading up to the event. If we think that history can span across time T, then the probability of the event occurring at any given time point t (t is an element of T) can be written as a probability density function $f(t)$. Then, the cumulative probability at time t is given by $F(t)$ and the integration of $f(t)$ from 0 to t. This is the probability of the event having occurred between time 0 and t. Then, the probability of survival, or the event not having occurred up to time t, can be simply written as $S(t) = 1 - F(t)$. Given $f(t)$ and $S(t)$, the hazard rate is the probability of the event happening at time t given the observation has not experienced the event until time t. With the simple notations outlined above, the hazard rate $h(t)$ is equal to $f(t)/S(t)$, the conditional probability of an event occurring, given that it has not happened up until time t.

Going back to the democratization example, the hazard rate at 10 years after a country's independence would be the probability of democratization at Year 10 divided by the probability of democratization not having happened for the past decade in this country. Theoretically, various factors would presumably influence the hazard rate. For instance, one may think that the level of economic development, former colonial experience, and strength of military force all affect the hazard rate, as well as the duration itself, and build the event history model accordingly. Event history analysis then estimates the hazard rate for democratization as the function of both these covariates and time. A baseline hazard rate refers to the hazard rate as only a function of time t.

Parametric and Nonparametric Approaches

There are various modeling options for researchers who seek to use event history analysis based on appropriateness to the data and background theories. First, the choice between parametric and nonparametric approaches depends on how confident researchers are of the shape of the baseline hazard, which ideally is guided by theory. In essence, all parametric models make explicit assumptions about the shape of the baseline hazard, while nonparametric models do not make such assumptions.

Parametric models make assumptions for the baseline hazard rate once covariates are included in the model. There are a wide variety of models, and models may be nested in others. For instance, the exponential model assumes that the baseline hazard is flat across time. This would mean that the probability of an event occurring at time t conditional on the event not having occurred is constant over time. The exact value then depends on included covariates. The Weibull model is more flexible than the exponential model, and it allows that the baseline hazard rate may be monotonically increasing, monotonically decreasing, or flat over time. Thus, the Weibull model nests the exponential model. Both the Weibull and exponential models assume proportional hazards as the changes of hazard rates with changes of covariates are proportional to the baseline hazard rate. The proportional hazards assumption should be tested. Since the Weibull model can be flexible, there have been many studies that have used it. Yet, in some settings, the monotonicity assumption may not be appropriate, and for those cases parametric models without a monotonicity assumption can be used. These models include the log-logistic and the lognormal models. These models allow hazard rates to first increase and then decrease as t passes. Neither of these models has the proportional hazards property.

The generalized gamma model can be useful to adjudicate among different parametric models, as several parametric models are nested within the generalized gamma model. The exponential, the Weibull, the lognormal, and the gamma models are all special forms of the generalized gamma model. When one has no a priori theoretical justification about how the baseline hazard rate varies across time, the generalized gamma model is likely to be particularly useful. If the fit is correct, parametric models generally have smaller standard errors than their semiparametric counterparts. However, as parametric models require assumptions about the shape of the baseline hazard, when assumptions are not correct, the estimation will be biased.

In comparison with the parametric models, nonparametric approaches do not make assumptions about the shape of the baseline hazard. While it is true that some parametric models are more flexible than others, they all still make assumptions about the shape of the baseline hazard. Moreover, more flexible models require more parameters to be estimated, and inclusion of such parameters may be tenuous and cumbersome. Finally, nonparametric models will approximate a correct parametric model; if the correct parametric model is the Weibull, nonparametric estimates will closely approximate the Weibull; if the correct parametric model is the exponential, nonparametric estimates will approximate the exponential without specifying the baseline hazard. In general, given that there is little theoretical justification for the shape of the baseline hazard in most political science inquiries, there seems to be a rising consensus that there is little to lose and much to gain in using nonparametric models in social science applications. Thus, it may be better to leave the baseline hazard unspecified and instead focus on how covariates of interest influence the hazard rate. This is the basic justification of the Cox proportional hazards model, or simply the Cox model. Time-varying covariates are easily handled in the Cox model as well.

The Cox model estimates the hazard rate of the ith observation at time t as a function of both the unspecified baseline hazard and the covariates. As the hazard rate is a product of the function of the covariates and the baseline hazard, the Cox model is also a proportional hazards model, and this assumption needs to be checked. The Cox model uses partial likelihood methods to estimate coefficients for each covariate. The partial likelihood function for some given data is the product of the hazard rate of the observation that experiences the nth event over the sum of all the hazard rates of observations that have not experienced the event, n to K, when n varies from 1 to N, where N is the number of observations that experience the event in the data, and K is the total number of observations. Then, the partial likelihood function is maximized, and the coefficients are obtained. For instance, imagine a data set with three observations ($K = 3$), where only two of them ($N = 2$) experience the event. The partial likelihood is then given by the hazard rate of the observation that experiences the event first over the sum of the hazard rates of all three observations multiplied by the hazard rate of the observation that experiences the event second over the sum of the hazard rates of the two remaining observations.

As illustrated, the Cox model uses information about which observation experiences the event sooner than others. Thus, the interval between two observations experiencing events does not provide any additional information. In this sense, the Cox model is an ordered events model. Important advances have been made in handling ties, which are observations that experience the event at the same time, such as the Breslow, the Efron, and the exact discrete approximations.

Discrete and Continuous Time Approaches

Thus far, the discussion of event history analysis implicitly assumes continuous time. While time is continuous in nature, in reality the data sets that researchers use are often discrete. Event history data in discrete time contain the same information in different forms—a series of binary observations in which zeros are recorded for a nonevent, and ones are recorded for an event. In such a case, researchers need to select a probability distribution to capture the binary decision. The logistic distribution (as in a logit model) and the standard normal distribution (as in a probit model) are typically used.

Duration dependence still needs to be directly accounted for in the discrete time model, which is generally handled by fitting smoothing functions such as spline functions or lowess. Some argue for discrete time due to greater familiarity by most social scientists with logit/probit models. Others argue for continuous time approaches since modeling duration dependency does not require extra modeling steps. Both discrete and continuous time models are appropriate and adequate.

Diagnostics

Model fit and assumptions should be tested for event history models, just as they are for any other statistical model. Residual plots show the difference between the observed values and the predicted values from the estimated model, and residual measures can be used as graphical tests for various purposes. Cox-Snell residuals are often used to evaluate the overall fit of the estimated model. When Cox-Snell residuals plot roughly around the 45° reference line, the Cox model performs reasonably well. When there is a systematic deviation of the plotted residuals from the reference line, then there may be some omitted variables or problems with model specifications. But one should also note that the Cox-Snell residuals are not a definite evidence of how the estimated model fits.

Since the Cox model and some parametric models assume proportional hazards, it is important to test this assumption. If it does not hold, it is not possible to interpret the hazard rate or ratio constant across time, and additional modeling steps are required. Graphically, scaled Schoenfeld residuals can be used in testing the proportional hazards assumption. Scaled Schoenfeld residuals can be thought of as the difference between the expected values of the covariates and the observed values of the covariates. Thus, when scaled Schoenfeld residuals are plotted across time, one can see if any variable shows violations of the proportional hazards assumption. When scaled Schoenfeld residuals are consistently gathered around the zero line, the proportional hazards assumption is likely to hold. When residuals systematically deviate from the zero line or show heterogeneity across time, it is possible that the model is misspecified and the proportional hazards assumption might not hold. There is a global statistical test to examine the proportional assumption as well. If the global test indicates that the model may have a violation, Harrell's rho can be used to focus on offending covariates individually. Typically, an interaction with time and the offending covariate is included in the model to account for the nonproportionality.

Extensions

Additional extensions make event history analysis even more useful. For instance, an observation may experience an event more than once. It is also possible that an observation can exit in more than one way. The model can also be extended to include unobserved heterogeneity among observations, event dependence, or spatial dependence.

Repeated Events

Repeated events are simply those when the observation can experience the event more than once, such as occurs in the study of repeated interstate disputes between certain pairs of countries. The analysis of repeated events takes into account the fact that the first, second, third, and such other events are not independent. Pairs of countries with two prior interstate disputes will likely have a different probability of having another dispute than pairs of countries with no such prior experiences of having an interstate dispute. Stratification by event number is an appropriate way to model repeated events to allow the hazard rate to vary by event number.

Competing Risks

Competing risks extensions to the basic Cox model allow analysts to consider different types of events. For example, a member of Congress may leave office due to defeat in the primary election, defeat in the general election, running for higher office, or retiring. Economic sanctions may end with either a sanctioning failure—a sanctioner lifting sanctions without achieving desired policy goals—or a sanctioning success—a target accepting the demand from the sanctioner. These events are distinct, and the covariate effects are likely to differ. The usual modeling strategy for a competing risks model is to estimate separate models for each event while treating observations experiencing other events as censored. One drawback of this approach is that it assumes independence between different types of events. Yet, often, this assumption is questionable. For instance, the hazard rate of sanctioning failure may depend on the hazard rate of sanctioning success. The competing risks model does not assume a time order for events to occur.

Unobserved Heterogeneity

In an ideal world, researchers would include all relevant covariates in the statistical model to estimate the precise effects of each covariate. But it is often not the case that all the relevant variables are available. Some may be unmeasurable, and others may be unobservable. The frailty model extension of the Cox model provides one way to account for unobserved heterogeneity across observations. The frailty term can be thought of as an additional factor that measures how prone or "frail" observations are to experiencing the event. The frailty can be modeled as either a group- or individual-specific term.

Event Dependence

The conditional frailty model separates and accounts for both event dependence and heterogeneity in repeated events models. Event dependence exists when the occurrence of one event makes further events more or less likely. For example, learning effects or damaging effects may make an event more or less likely to occur. In short, the risk of an event may be a function of a prior event occurring. Separating out event dependence and heterogeneity is helpful to analysts methodologically in order to draw more accurate substantive conclusions.

Spatial Dependence

Finally, recent extensions incorporate spatial dependence into event history models. This is an exciting development because it offers an effective approach to account for spatial dependence in political event processes. The model allows for spatially correlated random effects at neighboring locations. For example, scholars have looked at incorporating the proximity of civil unrest on the timing of outbreaks of additional violence and the diffusion of state policy adoption across states on the timing of those policy adoptions.

Janet M. Box-Steffensmeier
Ohio State University
Columbus, Ohio, United States

Byungwon Woo
Oakland University
Rochester, Michigan, United States

See also Censored and Truncated Data; Logit and Probit Analysis; Time-Series Analysis

Further Readings

Blossfeld, H., & Rohwer, G. (2002). *Techniques of event history modeling: New approaches to causal analysis* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Box-Steffensmeier, J. M., De Boef, S., & Joyce, K. (2007). Event dependence and heterogeneity in duration models: The conditional frailty model. *Political Analysis, 15*(3), 237–256.

Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event history modeling: A guide for social scientists*. Cambridge, UK: Cambridge University Press.

Darmofal, D. (2009). Bayesian spatial survival models for political event processes. *American Journal of Political Science, 53*(1), 241–257.

Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time-to-event data* (2nd ed.). Hoboken, NJ: Wiley-Interscience.

Kleinbaum, D. G., & Klein, M. (2005). *Survival analysis: A self-learning text* (2nd ed.). New York: Springer.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, UK: Oxford University Press.

Therneau, T., & Grambusch, P. (2000). *Modeling survival data: Extending the Cox model*. New York: Springer-Verlag.

Evidence-Based Policy

Evidence-based policy making (EBP) has developed over the past 40 years from conceptual roots that lie in the health care profession. Its deeper intellectual roots can be traced back even further to the Enlightenment, the policy sciences movement in the mid-20th century, and efforts to address societal ills through informed and activist government in the 1960s. Definitions of EBP may vary, but most observers agree that EBP requires rational and systematic processes for producing evidence to be applied in the policy making process. While differences in definitions of EBP tend to focus on variations on this idea, there is debate about the kinds of evidence that should be collected, EBP's feasibility as an actual guide for policy making, and the standards of evidence that it requires. In this sense, EBP shares challenges that are similar to earlier movements that sought to rationalize policy making and policy-making processes.

EBP's more recent conceptual roots lie in research on health care practices. In 1971, Archibald Cochrane, a British epidemiologist, highlighted gaps between the state of medical knowledge and the realities of medical practice. He advocated more systematic research on patient care as well as on the systematic application of knowledge gained through this research to clinical practice. In the early 1990s, his hope for more systematic assessments of clinical practices and their use in guiding patient care was realized when the international Cochrane Collaboration was established in Oxford, United Kingdom (UK), to execute systematic reviews of medical evidence. The work of the Cochrane Collaboration and the nearby Centre for Evidence-Based Medicine at John Radcliffe Hospital in Oxford caught the attention of policymakers in the UK, the United States, and elsewhere.

Research on evidence-based medicine gave rise to practical applications in both health care and other policy sectors. As Sue Dopson, Louise Locock, John Gabbay, Ewan Ferlie, and Louise Fitzgerald (2003) note, in the health care sector, the UK and the United States promoted the application of evidence-based research and practice in various ways, including the establishment of guidelines and research centers to encourage the use of systematic medical research findings in clinical practices. In the UK, the Labour government of the 1990s embraced the EBP concept and expanded it to other policy areas as a part of its efforts to modernize government. The Labour Government's 1999 Cabinet Office White Paper on Modernizing Government, for example, highlighted the importance of evidence and research in policy making across the board, not just in the health care arena. In an effort to ensure that public policies actually work to achieve their objectives, the UK also increased funding for social research during this time period.

In recent years, the phrase *evidence-based policy making* has been used with increasing frequency. For example, the title of the Association of Public Policy Analysis and Management's (APPAM) Annual Research Conference in Washington, D.C., during the fall of 2009, was "Evidence-Based Policy Making in the Post Bush/Clinton Era." This title reflects recognition of the importance of the phrase not only in the UK, where it has been commonly used, but also in North America and elsewhere as well.

However, while the EBP terminology has developed only in recent decades, the ideas underlying it have been in evidence for many years. Some trace the roots of EBP to the Enlightenment and its emphasis on applying science and knowledge for purposes of human betterment (Carolyn Heinrich, 2007). In the 1940s and 1950s, Harold Lasswell's efforts to build a science of policy through policy analysis sought to encourage analytically based policy improvement efforts that are similar to those targeted by EBP. And, in the 1960s, governmental efforts to create a "Great Society" in the United States were predicated on the idea that rationally constructed public policies could be implemented in ways that alleviate societal ills. These policies, in turn, were accompanied by the establishment of public policy think tanks in universities and elsewhere that were designed to provide objective information to help guide efforts to improve policy making on an ongoing basis.

The attention that EBP has received is in part traceable to its grounding in basic concerns about the application of human knowledge to improve public policies, but it also has roots in the unique circumstances of the past several decades. During the 1980s, growing skepticism regarding the effectiveness of government programs led to a search for new alternatives that took better account of their costs and the cost-effectiveness. These concerns contributed to the performance measurement movement of the 1990s and a need for government programs that "work better and cost less," to borrow a phrase from the famous 1993 Gore report in the United States. The EBP movement fit in nicely with these concerns and added a specific focus on the role of research and evidence to the performance measurement interests that came to dominate public sector management in the 1990s and the early years of the 21st century.

Challenges for Evidence-Based Policy

However, while the EBP movement coincided well with the times in which it developed, it did not escape the issues that had plagued similar movements in the past. At least three issues are worth addressing in this regard.

First, the concept of EBP assumes that researchers and/or policymakers can define what evidence is important and apply it appropriately to achieve desired policy outcomes. However, public sector programs may seek to achieve more than one goal, and choices regarding the most appropriate evidence to collect are not always clear. The evidence one needs to garner depends on the goals one chooses to pursue or assess. Where goals are numerous and differ with respect to the availability of evidence and the ease with which it can be developed, the evidence produced may be inherently biased toward goals that are easily measured. As a result, some argue that EBP and performance measurement more broadly are inherently biased toward measures of efficiency that are more easily quantified in monetary terms than measures of other values, such as those relating to equity, fairness, and democratic participation. This concern is exacerbated by the fact that producing evidence often requires money, which may be available in less abundance to those in society who are most concerned about these other values.

Second, there are debates about whether evidence produced by research efforts can be incorporated effectively into policy-making processes. Research takes time to be done in quality fashion, and time is often in short supply as policymakers make decisions to address both perceived emergencies and ambitious policy agendas. There are also issues relating to the ability of policymakers to access the right or best information as they make decisions. Because policymakers are often generalists, they may have limited ability to grasp all aspects of the evidence they receive. As a result, they may depend heavily on those who provide them with evidence, and they may ultimately rely more on the perceptions of those who provide the evidence than on the evidence itself—particularly when the evidence is complex and subject to interpretation. Finally, because policymakers in democratic political systems are (at least theoretically) held accountable to their electorates, they may perceive a need to react to public perceptions of evidence and risk rather than to the evidence itself. And where public perceptions of risk and evidence differ from the actual evidence provided by experts, systems of democratic accountability may encourage decisions based on public perceptions rather than on objective evidence.

A third difficulty that is inherent in EBP relates to decisions regarding the adequacy of evidence that is provided. Even within scientific disciplines, there is often disagreement among experts regarding the kinds of evidence that are most appropriate for answering particular kinds of questions. This is true in the social sciences as well as in the natural sciences, and it is particularly true when evidence is being gathered to help guide policy making because of varying values that may come into play in these cases. Some researchers argue that quantitative information is critically important if one is to assess policy effectiveness in objective fashion, while others argue that more qualitative assessments are also required if one is to reach balanced conclusions that address a range of legitimately held values in society. As a result, even where evidence is timely, appropriate, and clear to all involved, there are likely to be cases where differing kinds of evidence lead one to differing conclusions. And there is certainly room for debate over what kinds of evidence are most appropriate to believe in these contexts.

In all three of the aforementioned areas, EBP faces challenges that parallel the challenges faced by similar movements in the past. While EBP has the benefit of past experiences and appears to be particularly cognizant of these challenges as a result, it is still working to address these questions. What kind of evidence should be collected? How should evidence be incorporated into policy-making processes? And what standards of evidential quality should be applied? These are not simple questions, but they are important ones. There is reason to be skeptical of the extent to which these questions will be answered adequately once and for all, but—if one believes that human intelligence and the rigors of systematic inquiry can yield beneficial impacts—there appears to be little alternative to addressing them in some fashion. In the end, EBP and its predecessor movements focus on bringing systematic evidence to bear on the policy-making process and—despite the questions raised above—this may be the best kind of process that is available.

John A. Hoornbeek
Kent State University
Kent, Ohio, United States

See also Policy Analysis; Policy Evaluation; Policy Formulation; Policy Framing; Policy Instruments; Policy Network; Policy Process, Models of

Further Readings

Cochrane, A. L. (1971). *Effectiveness and efficiency: Random reflections on health services.* Cambridge, UK: Royal Society of Medicine Press.

Davies, H., Nutley, S., & Smith, P. (2000). *What works? Evidence-based policy and practice in public services.* Bristol, UK: Policy Press.

Dopson, S., Locock, L., Gabbay, J., Ferlie, E., & Fitzgerald, L. (2003). Evidence-based medicine and the implementation gap. *Health, 7,* 311–330.

Gore, A. (1993). *Creating a government that works better and costs less: The report of the national performance review.* New York: Plume.

Heinrich, C. J. (2007). Evidence-based policy and performance management: Challenges and prospects in two parallel movements. *American Review of Public Administration, 37,* 255.

Nutley, S. (2003, April 23–24). *Bridging the policy/research divide: Reflection and lessons from the UK.* Paper presented at National Institute of Governance Conference, "Facing the Future: Engaging Stakeholders and Citizens in Developing Public Policies," Canberra, ACT, Australia. Retrieved November 18, 2010, from http://www.treasury.govt.nz/publications/media-speeches/guestlectures/nutley-apr03

EXECUTIVE

The national executive—more usually referred to as the "government" or, in the United States, as the "administration"—is the only political institution that exists in every country. There cannot be a state without a national executive, while bodies such as legislatures, parties, or interest groups exist only where the polity is, at least relatively speaking, liberal-democratic. This entry concentrates exclusively on liberal-democratic polities, which have become, by the early years of the 21st century, a substantial majority of the 200 or so polities of the world; yet there are still countries without legislatures, without parties, or even without interest groups, either because they are very traditional, as some of the Middle Eastern states, or because a coup, typically of a military character, replaced a liberal-democratic political system that had been in place for only a few years.

Both the composition and the power of national executives need to be examined with care—the composition because it varies appreciably in terms of the relationship among government members and because of the role of the leader of the government, and the power because of the variations and indeed even marked disagreements over what governments can effectively achieve with whatever powers they may formally have. This entry therefore looks successively at both aspects; before doing so, however, it briefly describes the way in which national executives gradually developed.

Evolution of the National Executive

While the proportion of liberal-democratic national executives is large in the contemporary world, as noted, this was not the case at all in the past, even in much of the 20th century, let alone earlier. It is worth noting that, in the classical Roman Republic, there were a number of separate executive positions rather than a government—for instance, the positions of consul, pretor, or edile. A similar framework was adopted by the few republics that came to exist during the Middle Ages and later, principally in Italy. The idea of a government prevailed in monarchies, as kings or queens came to appoint "secretaries" or "ministers" to oversee the fields that had to be covered, such as those of finance, foreign affairs, or defense. This model was widely adopted in the strong—or absolute—monarchies of 16th-, 17th-, and 18th-century Europe. As against this trend, the first modern—and liberal—republic, that of the United States, was established by a formal constitution that came into force in 1789 and proved to be a great success; that is, it endured and was imitated early in the 19th century, as a variety of Latin American states became independent from Spanish rule.

On the other hand, the French revolutionary republic, set up in 1792, was not a success: It was replaced after a few years by Napoleon's empire, and the monarchy was reestablished in 1815. Yet that monarchy was different in character from the one that had ruled France for centuries; this was the result of the profound changes in the direction of liberal government that had begun to occur in Britain over a century earlier, from 1688. Such a move toward liberal government in Britain was followed gradually after 1815 through the course of the 19th century in parts of Western Europe and beyond, including Japan. The evolution toward a liberal-democratic government was to lead to the emergence of a parliamentary-cabinet system, which was to characterize much of Western Europe and the "Old Commonwealth" countries from the end of the 19th century in contrast to the American presidential system.

The move toward liberal democracy in Europe was interrupted between the two World Wars, however, as the liberal-democratic model was shaken and indeed replaced by dictatorships. Revolutionary Russia, renamed the Soviet Union, became the first Communist state from the 1920s to the 1990s; in the 1920s and 1930s, right-wing dictatorships were installed, by coups or even as a result of elections, in many parts of continental Europe, especially in Italy, Germany, and Spain, as well as in most of the states that had been created after World War I in East-Central and Eastern Europe. Moreover, although World War II ended with the victory of democracies in 1945, only in the 1980s and 1990s did the trend toward liberal democracy become widespread. Many of the new states of Asia and Africa had indeed quickly become dictatorships after having acquired independence from European colonial powers in the 1950s, 1960s, and 1970s; the major exception being India that was to remain, from 1948 onward, by far the most populous liberal democracy.

Structure and Composition of the National Executive

The national executive is the body that governs the state, both in democratic and in nondemocratic countries; exactly who composes that executive is not always entirely clear, and the relationship among its members is both varied and sometimes rather obscure. As was already suggested, liberal-democratic executives are often said to be divided into two types: the presidential and the parliamentary-cabinet type. While that dichotomy does correspond to a major distinction, there are variations within each of the two groups. There is also an intermediate type known as semipresidential, singled out and theoretically developed by Maurice Duverger (1980) and Robert Elgie (1999).

The "ideal-type" presidential model is based on a sharp division among three state powers—executive, legislative, and judiciary—the assumption being that none of these powers can encroach on the others, a distinction that is particularly relevant with respect to the relationship between the executive and legislative powers. The model originated in the United States at the time of independence; while the American constitution of 1787 (ratified in 1789) has been amended many times, it has kept its main original features. The cornerstone of the executive is naturally the presidency, whose holder is elected for 4 years by the people (albeit indirectly, as electors vote for the members of an electoral college that has as its only function the election of the president). The Twenty-Second Amendment (1947) further mandated that the president can be reelected only once, while in some Latin American states (most of which adopted a similar model) the president cannot serve two successive terms and in some cases cannot be reelected at all.

The U.S. president is free to appoint the members of the administration at will, except that the upper house of the legislature—the Senate—has to approve these appointments. The top echelon of the executive is constituted by what is formally described as a cabinet composed of secretaries in charge of the various departments of the government, under whom are assistant secretaries, also appointed by the president. There is also an Executive Office of the President, while the closest advisers are part of the White House Office. All these are appointed at the "pleasure" of the president and rarely remain in their position from one presidential administration to the next, even if the incoming and outgoing presidents are members of the same party.

Perhaps the key characteristic of the American executive (and probably of at least many of those presidential systems modeled on the American system) is that its members are only loosely connected together and often do not know each other well at all; indeed, the president may not have had a close relationship with at least a number of them, although some of the positions may be filled by friends who are rewarded for their help, in particular during the presidential election campaign. This does create a situation that Hugh Heclo (1977) has described as a "government of strangers." It also follows that one does not really know precisely who is part of the government. Problems are particularly serious with respect to the distinction between the secretaries and their assistants, on the one hand, and White House staff members, on the other, the latter being often effectively in charge of some of the functions of the secretaries, particularly in the context of foreign affairs, as when Henry Kissinger was adviser to President Richard Nixon before he became Secretary of State in the early 1970s.

The structure and composition of the national executive are very different in the other models of liberal-democratic executives; for example, that of the parliamentary-cabinet system, which, following British developments since the late 17th century, has been adopted widely in Europe and in many Commonwealth countries, as well as, among others, in Japan, Thailand, Israel, and Turkey. Originally, as was suggested earlier, the system allowed the members of the government—typically a cabinet—to run the affairs of the state, at least nominally, on behalf of the monarch.

The cabinet system differs in two fundamental ways from the presidential model, first in that the cabinet must enjoy the "confidence" of parliament to be able to be and remain in office and, second, in that the cabinet takes its decisions "collegially": The prime minister is the head of the cabinet and therefore a part of it, while American presidents are not strictly speaking a part of their cabinet. As a result, members of the cabinet form a team: They stand or fall together; moreover, given the part played by parliament in the whole structure, cabinet members are typically drawn from among members of parliament. A further indirect consequence is that the duration of a cabinet is not fixed in advance but depends on the maintenance of parliamentary support.

Furthermore, and again largely because the cabinet needs the support of a majority in parliament, cabinets may be—and often are—based on coalitions among two or more parties. There can be single-party governments when one party enjoys an absolute majority in parliament or when, as occurs fairly frequently in Scandinavian countries, the government commands minority party support only but has negotiated (typically in advance) a promise from some of the other parties that they will vote for the impending government. In this case, but also in cases of coalitions, formal agreements that aim at determining in advance what the policy of the government will be are needed.

While the president is the head of the executive that he or she has set up, in parliamentary government the prime minister does not have, as a result of his or her position alone, a formally dominant position in the parliamentary-cabinet executive. This may happen, but many situations prevent the prime minister from exercising the kind of supreme leadership that the president is often said to have. Yet the president is also constrained in many ways, despite the fact that the expression *presidentialization* is sometimes used to refer to very powerful prime ministers.

Because the presidential system is regarded as somewhat rigid (Juan Linz, 1994), an intermediate system, known as semipresidentialism, has been adopted in some countries, France in particular, which is where it has lasted the longest, in effect for half a century since 1958; but it has also prevailed, in a less clear-cut manner, admittedly, in Finland and in Portugal, as well as outside Europe, in Sri Lanka, Taiwan, and South Korea. The detailed organization and practice of semipresidential systems varies appreciably, but the general principle is that, alongside the president, who is popularly elected, a cabinet headed by a prime minister is, as in parliamentary-cabinet systems, responsible to parliament. Finally, to be truly comprehensive, the case of Switzerland must be mentioned, since this is the only country in which the national executive, known as the Federal Council, which is composed of seven members elected for 4 years by parliament, is truly collective.

To be able to compare effectively these forms of government, and in particular presidential and parliamentary-cabinet systems, this entry turns to an examination of the power of these bodies, both formally and in practice, to direct the policies of the nations at the head of which they are.

Power of National Executives

Executive power is difficult to define. It is not based only on the powers that, for example, constitutions give to these bodies. As a matter of fact, most constitutions are somewhat vague or even misleading in this respect, to the extent that they may suggest that legislatures are in charge of the lawmaking process. This may be true from a purely constitutional point of view, but it is rarely true in practice. Moreover, the word *executives* conveys the impression that executives merely execute policies or ideas that have originated elsewhere. In reality, both in parliamentary-cabinet systems and, albeit to a more limited extent, in liberal-democratic presidential systems, legislation is prepared by the government.

Executives have in reality to fulfill three different functions: those of initiation, coordination, and implementation. Initiation means that the ideas that may have been put forward previously by parties, especially by those that are represented in the government, as well as by interest groups or nongovernmental organizations (NGOs) and indeed by members of the legislature have to be implementable, both administratively and politically. They must not arouse conflicts such that the ability of the government to achieve its overall goals will be affected. Second, executives must coordinate the policies that have been initiated, as some of these may be closely connected to others; moreover, the policy-making process as a whole has to be financially sustainable: Proposals may have to be postponed, despite the fact that they may have previously been announced as being part of the government program. Finally, policies have to be implemented. It is often believed that implementation is simple and can be left to subordinates. This is true only up to a point, as implementation may be difficult, for instance, if a new organization has to be set up or if a matter is regarded as particularly urgent.

The skills required of ministers to fulfill satisfactorily these three governmental functions are very different. Initiation implies imagination to ensure that the new policy will be attractive and effective, coordination entails being able to work closely together with other members of the government, and implementation means managing public service bodies that are often very large and whose permanent officials may prefer to postpone new developments. Many members of the executive may find it difficult to carry out all three sets of activities well. The problem is particularly serious in parliamentary-cabinet systems as these are officially collective and at least collegial. It is in this type of executive that the members of the cabinet—the ministers—are expected to have the skills required to initiate, coordinate, and implement policies.

The following discussion first examines the case of members of the executive in presidential systems and, specifically, of liberal-democratic presidential systems, such as the American polity. As was noted earlier, there is a cabinet in the United States, but that cabinet is not in charge of the overall policy making; the president is in charge. The members of the cabinet—the secretaries—are essentially regarded as managers of policies that have been previously adopted in principle by the president: They may have helped frame these policies, together with their officials of the federal civil service, but the decision to proceed has to be given by the president. Coordination is also likely to be the province of the Executive Office of the President, where the Office of Management and Budget is located.

Yet each president tends to hold different views about how the whole executive is to be organized as well as about the ways in which the various offices are to relate to each other; moreover, as we noted, the secretaries are likely to be strangers to each other and to the departments that they run. As Shirley Warshaw (1996) has noted, there have therefore been substantial variations from president to president about the process of policy making, and the results have often been unsatisfactory.

These kinds of difficulties seem to be in the nature of the liberal-democratic presidential system, especially when that executive is very large, as it is in the case of the American federal government; yet two further sets of problems arise almost inevitably. The first has to do with the fact that there is clearly some ambiguity about the effective power of the executive with respect to the initiation of policies, especially when these require new legislation. As in liberal-democratic presidential systems, the legislature is elected independently from the executive, and it tends to maintain substantial control over the proposed bills that are submitted, including those that the president sends; it also exercises much independent control over the budget, in terms of both income and expenditure. There is therefore a clear limit to the influence that presidents can exercise. They may cajole and induce; they cannot threaten.

Second, at the other end of the chain, there is also a limit, in practice, if not formally, to what the president and the secretaries can extract, so to speak, from the departments. Most departments cherish some kinds of policy developments more than others; meanwhile, they cultivate close relationships with some members and in particular some committees of the Congress; this enables the officials of these departments to obtain financial support for their requests. As a result, the top members of the executive—the secretaries—are often likely to find it difficult to change the direction of the policies of their departments. In the late 1960s and early 1970s, for instance, President Lyndon Johnson, who was a policy expansionist, and President Richard Nixon, who was in favor of retrenchment, experienced problems in convincing departments to achieve their own goals in view of what can be regarded as the policy "habits" of the departments of the federal government.

Thus, in practice, there are clear limitations to the power that the president and the executive in general can exercise in a liberal-democratic presidential system such as that of the United States. Such limitations do not occur to the same extent in those presidential systems in which the presidential powers are less constrained by the constitution, as stressed by Mathew Shugart and John Carey (1992). Most political scientists believe that there were fewer limitations to the effective power of the executive, seemingly, in the few liberal-democratic semipresidential systems that existed in the beginning of the 21st century, of which France is often regarded as the best example. Finally, the executive appears to be least constrained of all in parliamentary-cabinet systems principally because that executive, more than the American executive, usually controls the legislature.

In parliamentary-cabinet systems, the government does indeed need majority support of the legislature to remain in existence and, by and large, that support exists because it is sustained by one or more parties that constitute the majority. According to Jean Blondel and Maurizio Cotta (2000), this is why that type of executive has often been described as being based on party government. Since there is a close relationship between the executive and the legislature in the parliamentary-cabinet system, the cabinet is able to exercise enough pressure on the legislature to obtain support at any rate for a very large proportion of the bills that it introduces in parliament as well as for a very large proportion of the financial requirements that it submits for approval. This does not mean that parliament can never have its way on any issue or on any budgetary provisions; there are indeed differences from country to country about the extent to which parliament can induce the executive to modify its policies, but by and large, across Western and even East-Central Europe, as well as in those countries of the Commonwealth that are liberal-democratic parliamentary systems, the executive is broadly speaking able to achieve its goals. If this is not achieved, there will be a governmental "crisis": The prime minister is likely at least to threaten to resign and, indeed, in some cases, the threat is implemented.

The power of the executive in parliamentary cabinets of a liberal-democratic kind is thus large; yet it does not follow that all the members of these cabinets, let alone all the members of the executive, including those who are below cabinet rank, can exercise an equal fraction of that overall power. It has been pointed out by Michael Laver and Kenneth Shepsle (1996) that ministers are often more concerned about what goes on in the department that they run than about overall governmental policy. Although all members of the cabinet formally share the overall power of initiation that was referred to earlier, many do not exercise that power in fact at cabinet meetings. In parliamentary-cabinet systems, these meetings take place frequently (typically, once a week); yet they would be overburdened if every one of the 15 to 20 members of the cabinet were to participate actively on every issue. The result is that, by and large, the prime minister exercises more than the usually recognized role of "first among equals" and indeed that some prime ministers, in Britain but also elsewhere, are regarded as being "presidential," wrongly in a sense because, as was noted earlier, presidents of liberal-democratic systems are confronted with the substantial power of the legislature and with the tendency of departments to be somewhat autonomous in the way they conduct their affairs. The power and strength of prime ministers do indeed vary appreciably, partly across countries, but partly also as a result of the extent to which they wish to be activists, in leading, indeed possibly in forcing, members of their cabinet to follow their line (Brian Farrell, 1971; Anthony King, 1994). It is nonetheless usually recognized that prime ministers of parliamentary cabinets enjoy a degree of power over political developments in their countries that exceeds that of presidents of liberal-democratic systems.

Conclusion

Executives in the contemporary world are, by and large, very powerful, even if they head liberal-democratic systems in which there are constitutional and customary limitations to what governments can do—that is to say, their members are constrained by "veto players," as George Tsebelis (2002) called them. Admittedly, some decisions may be slow to take because of the influence that interest groups and NGOs are able to exert and of the consequential unwillingness of governments to attempt, or attempt quickly, to override the proposals of these bodies.

Yet it is not the case that liberal-democratic governments cannot achieve what they propose to do, especially if these proposals are part of the program that was presented to the electorate at the previous election. While presidents in presidential systems of the liberal-democratic variety are typically, especially in the American case, not closely constrained by these election programs, parliamentary-cabinet governments are; this constraint limits effectively the power that these executives would be able to exercise if they were not party governments. There is thus little danger that governments should become unable to act because of the constitutional or customary limitations under which they operate. The limitations to which they have to pay strong attention stem from the part that the civil society at large is able to take in a democratic society, and it is surely right that it should be so if liberal-democratic systems are to deserve their name.

Jean Blondel
European University Institute
San Domenico di Fiesole, Italy

See also Cabinets; Democracy, Types of; Government; Leadership; Parliamentary Systems; Presidentialism; Semipresidentialism

Further Readings

Blondel, J. (1982). *The organisation of governments*. London: Sage.

Blondel, J., & Cotta, M. (Eds.). (2000). *The nature of party government*. Basingstoke, UK: Palgrave.

Blondel, J., & Muller-Rommel, F. (Eds.). (1988). *Cabinets in Western Europe*. London: Macmillan.

Duverger M. (1980). A new political system model: Semi-presidential government. *European Journal of Political Research, 8*(2), 165–187.

Elgie, R. (1999). *Semi-presidentialism in Europe*. Oxford, UK: Oxford University Press.

Farrell, B. (1971). *Chairman or chief? The role of the Taoiseach in Irish government*. Dublin, Ireland: Gill & Macmillan.

Heclo, H. (1977). *A government of strangers*. Washington, DC: Brookings Institution Press.

King, A. (1994). Chief executives in Western Europe. In I. Budge & D. McKay (Eds.), *Developing democracy* (pp. 151–163). London: Sage.

Laver, M., & Shepsle, K. (1996). *Making and breaking governments*. Cambridge, UK: Cambridge University Press.

Linz, J. J. (1994). *The failure of presidential democracy*. Baltimore: Johns Hopkins University Press.

Shugart, M. S., & Carey, J. M. (1992). *Presidents and assemblies*. Cambridge, UK: Cambridge University Press.

Tsebelis, G. (2002). *Veto players*. Princeton, NJ: Princeton University Press.

Warshaw, S. A. (1996). *Powersharing: White House cabinet relationships in the modern presidency*. Albany: State University of New York Press.

Experiments, Field

Field experimentation brings the rigorous estimation of causal effects that is the hallmark of experimentation to more naturalistic settings. This entry begins by defining field experimentation. It then describes the early history of political science field experimentation and discusses more recent developments. It concludes with a discussion of some of the weaknesses of field experimentation.

Definition

Experimentation is a method for obtaining unbiased estimates of causal effects. In social science experiments, the unit of observation is randomly assigned to different interventions, and the effect of the intervention is measured by comparing outcomes of interest across the randomly assigned groups. Field experiments are experiments that take place in real-world settings and attempt to reproduce the environment in which the phenomenon of interest naturally occurs. In contrast, laboratory experiments are often highly stylized representations of the social behaviors the investigators wish to learn about through their experiments, leading to concern that the results from the laboratory will not apply to behavior more generally. The realism of field experiments is intended to minimize concerns about the external validity, or the generalizability, of the experimental results.

Experiments have many components, which may differ in their degree of realism, leading to a blurring of the distinction between what is or is not a field experiment. Economists Glenn Harrison and John List propose a four-category taxonomy to distinguish field experiments from laboratory experiments. According to this taxonomy, a conventional laboratory experiment uses a standard subject pool (i.e., drawn from the university) and an abstract framing of the problem being investigated. Examples of this would be laboratory studies of play in dictator and ultimatum games. Two intermediate categories of experiments, which move in the direction of greater realism, are the artifactual and the framed field experiments. An artifactual field experiment deviates from the conventional laboratory experiment through the use of nonstandard subjects. For example, researchers have investigated ethnic conflict by performing standard laboratory games, such as the dictator game, in Africa, with subjects drawn from groups with a history of ethnic rivalry. The framed field experiment moves further away from the conventional laboratory experiment. This type of experiment uses realistic subjects (as does the artifactual experiment); however, in contrast to the artifactual experiment, the task and context is also more realistic. An example of a framed field experiment in political science is research on the effect of money on legislative access, in which congressional staffers make hypothetical scheduling decisions based on information they are provided about the individual seeking time with the representative. A natural field experiment, which is the design often referred to simply as a "field experiment" in political science, is the same as a framed field experiment, except that the environment is the one where the subjects naturally perform the task in question and the subjects are unaware that they are in an experiment. Research in which political campaigns randomly assign households to receive different campaign mailings to test the effect of alternative communications on voter turnout is an example of a natural field experiment.

This entry focuses on natural field experiments. Although the degree of naturalism in field experiments is the greatest strength of the method, it is important to remember that the goal of experimental interventions is to estimate a causal effect, not to achieve "realism." If the researcher aims to capture basic psychological processes that do not vary across populations, experimental contexts, or subject awareness of the experiment, then there is no problem with conventional laboratory studies. That said, understanding behavior in natural environments is frequently the ultimate goal of social science research, and it is hard if not impossible to even recognize the full set of threats to external validity present in artificial contexts, let alone adjust the measured experimental effects and measures of uncertainty to account for these threats.

Field Experiments in Political Science

The earliest field experiments in political science were performed in the 1920s by Harold Gosnell, who sought to measure the effect of get-out-the-vote appeals in the 1924 presidential and 1925 mayoral elections in Chicago. Over the next 30 years, researchers conducted related field experimental work on the relative effectiveness of emotional versus rational political appeals and the turnout effects of different modes of communication (mail, phone, and visits). This early literature appears to have been inspired by the development and application of experimental methodology in agriculture, medicine, and other fields. These pioneering field experiments, however, had limited impact on subsequent research strategies. With rare exceptions, there was no follow-up work using the experimental method in real-world settings to replicate the earlier experiments, to extend the voter mobilization studies to new settings or communications methods, or to apply the method to new questions. Rather, when considering alternative research designs, field experimentation was either ignored or dismissed due to skepticism about whether real-world experiments were feasible. Indeed, no field experiment appeared in a major political science journal in the 1990s.

In recent years, there has been a significant revival of field experimentation in several disciplines, most notably in economics and political science. This turn to field experimentation follows important intellectual developments across the social sciences. During the 1980s, there was increasing appreciation, especially in economics, of the difficulty in estimating causal effects from standard observational data due to bias caused by unobservable factors. Researchers were increasingly attracted to naturally occurring randomizations or near-randomized applications of a "treatment" that could provide data where, as in a planned experiment, there is no expected association between the treatment and the unobservables. Examples of this work include Joshua Angrist's study of the effect of Vietnam War service on wages, where an individual's draft lottery number altered the likelihood of service, and Steven Levitt's use of variation in policing levels associated with election cycles to estimate the effect of police on crime. This brand of research was characterized by enormous creativity. However, naturally occurring randomizations or near randomizations are rare, perhaps as rare as the talent required to spot them, and valid instrumental variables must satisfy rigorous requirements. In contrast, randomized experimental interventions permit scholars to produce exogenous variation in the treatment variable by design, avoiding the need to discover a fortuitous naturally occurring randomization or make the often questionable assumptions required by other statistical designs. A further advantage of planned experiments is that the researcher knows which subjects are assigned to be treated. This is frequently not the case in observational studies. In particular, for studies based on survey data, treatment assignment is typically measured through self-reports, introducing the possibility of bias due to measurement error.

The modern revival of field experiments in political science can be traced to Alan Gerber and Donald Green, who in 2000 examined the effect of voter mobilization in New Haven, Connecticut, on the November 1998 elections. Gerber and Green measured the turnout effects of nonpartisan face-to-face canvassing, phone calls, and mailings. Since this study, numerous field experiments have measured the effects of political communications on voter turnout. Several replicate the basic design of the New Haven study and measure the effect of canvassing, phone, or mail in new political contexts, including other countries. Others examine new modes of communication or variations on the simple programs used in New Haven, such as an analysis of the effect of phone calls or contacts by communicators matched to the ethnicity of the household, repeat phone calls, or partisan television and radio broadcasts. Recent work has begun to consider the effects of newer technologies, such as e-mail and text messaging, or novel approaches to mobilization, such as election day parties at polling places. Meta-analysis of the results of dozens of canvassing, mail, and phone studies is reported in a quadrennial literature review, "Get Out the Vote." Overall, the results have been consistent with the initial New Haven findings. Canvassing has a much larger effect than less personal modes of communication, such as phone and mail. The effect of brief commercial calls, such as those studied in New Haven, and nonpartisan mailings appear to be less than 1 percentage point while canvassing boosts turnout by about 7 percentage points in a typical election.

In recent years, field experimentation has spread beyond measuring the effects of various interventions on voter turnout to address other aspects of political behavior, to assess institutional performance, and to consider broader theoretical issues such as social influence, norm compliance, collective action, and interpersonal influence. For instance, one recent application of field experiments to investigate theoretical constructs in political behavior uses a field experiment to measure the causal effect of partisanship. Researchers sent mailings informing unaffiliated, registered voters of the need to affiliate with a party to participate in the upcoming closed primary in Connecticut. They found that the mailings increased formal party affiliation, and a posttreatment survey found a shift in partisan identification as well as a shift in political attitudes. Another study measures social spillover effects of political communications and finds that spouses and roommates of those who are contacted during a voter mobilization drive are also more likely to vote. Other scholars have investigated the effect of social pressure on norm compliance by measuring the effects of alternative mailings that exert varying degrees of social pressure. These findings indicate that a pre-election mailing listing the recipient's own voting record and a mailing listing the voting record of the recipient and their neighbors cause a dramatic increase in turnout. Follow-up research uses a postexperiment survey to measure whether this treatment response varies with subject personality. A novel study of elite behavior examines the effect of a lobbying effort on a bill in the New Hampshire legislature and finds that a random e-mail from an interest group causes a statistically significant increase in roll call voting for the sponsor's measure.

Although the initial studies in the modern wave of field experiments were entirely by American politics specialists, comparative politics and international relations scholars are now applying the method to important questions in these fields. For example, one study compares broad policy versus narrow clientelistic campaign messages in a 2001 Benin election. Another study examines the effect of election monitors on vote fraud levels. In a study of collective action and social capital conducted following civil war in Liberia, researchers found that foreign aid for community-driven reconstruction projects increases social cohesion as measured by behavior in public goods games. A recent example of work that examines institutional performance compares the performance of alternative institutions for the selection of a public good in Indonesia. This research finds that while more participatory institutions do not change the set of projects approved, participants are more satisfied with the decision-making process. A recent study examined the effect of legislative performance report cards on representatives' attendance records in Uganda. They report that instituting this method of showing legislators' attendance to their constituents results in higher levels of parliamentary attendance rates. This work highlights an important contrast between field experimentation and observational methods: Field experiments may be used to study both common real-world phenomena (such as campaign television commercials or the effect of election monitors) as well as novel interventions for which there are no observational counterparts (unusual mailings or legislative report cards in developing countries). For novel interventions, there are no naturally occurring parallels to the intervention, and so no observational study is even possible.

In addition to these and other substantive applications, field experiments have increasingly been used to assess the performance of nonexperimental methods. Following the example of Robert LaLonde's study of job training programs, data from experimental studies can be reanalyzed using observational techniques. The performance of the observational method can then be evaluated by comparing the results with the unbiased estimates obtained when the analyst uses knowledge of the randomized treatment assignment. Applying this approach to an experimental study of voter mobilization phone calls, researchers compared the experimental and observational estimates of the effect of the intervention. They find that regression analysis and exact matching both dramatically overestimate the treatment effect.

Some Weaknesses of Field Experimentation

The combination of unbiased estimation and a natural context makes field experimentation very attractive, but the method has some important drawbacks. Among the various criticisms, three are focused on here. First, field experiments are typically more costly and difficult to implement than conventional laboratory experiments or standard observational studies. A field experiment designed to test the effects of face-to-face canvassing involves recruiting and training canvassers and then sending out canvassers to deliver mobilization messages, perhaps in the rain, to subjects who may or may not be home. Following the election, the data must be merged with the postelection voter file. It is logistically more challenging than inviting subjects into your university laboratory to listen to a script and then state their vote intention. Scholars have analyzed the design of experiments to improve their statistical efficiency, an especially important concern for field experiments facing low compliance rates.

Second, in contrast to the laboratory environment, interaction between subjects cannot be controlled in the field environment, which leads to the possibility that the outcomes for subjects will be influenced by the treatment status of other subjects. For example, a study comparing treatment group and control group households will be biased if those who are treated interact with those assigned to the control group and in turn alter the behavior of those in untreated households. These contamination effects can be minimized by measuring the outcome soon after treatment or treating a small fraction of the subject pool. Furthermore, contamination effects can themselves become the object of study, either by asking about interactions across subjects in a posttreatment survey or by measuring the effects of the proximity of control group subjects to those randomly treated. These spillover effects are of course endemic in the data routinely used in observational studies, though this complication is rarely noted.

Third, the set of field experiments that make it into print may suffer from selection bias. Publication bias is a general danger, but there are several reasons for special concern regarding field experiments. To offset the costs of field experiments, researchers may choose to collaborate with nonacademic organizations or groups that undertake the activities being studied. This poses two dangers. First, the subset of organizations wishing to participate and, therefore, the results of a study of their activities may be unrepresentative of all active organizations. Bias will follow when the willingness to be studied is a function of the expected efficacy of the group's efforts. Second, and perhaps more significantly, after the experiment has been performed and analyzed, organizations may wish to keep the results as trade secrets or to suppress negative findings. A final, and related concern, is that field experiments typically require more effort to replicate than the usual laboratory experiment. In principle, when a researcher reads an interesting social psychology paper, he or she can easily attempt to reproduce the same finding in his or her psychology lab at a small cost and with little effort. In contrast, if field experiments appear only intermittently, it may be some years before a misleading finding is corrected by the accumulation of contrary evidence.

Alan S. Gerber
Yale University
New Haven, Connecticut, United States

See also Experiments, Laboratory; Statistical Inference, Classical and Bayesian

Further Readings

Davenport, T. C., Gerber, A. S., & Green, D. P. (2010). Field experiments and the study of political behavior. In J. E. Leighley (Ed.), *The Oxford handbook of American elections and political behavior* (pp. 69–88). New York: Oxford University Press.

Green, D. P., & Gerber, A. S. (2008). *Get out the vote: How to increase voter turnout* (2nd ed.). Washington, DC: Brookings Institution Press.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature, 42,* 1009–1055.

Humphreys, M., & Weinstein, J. M. (2007, August). *Policing politicians: Citizen empowerment and political accountability in Africa.* Paper presented at the annual meeting of the American Political Science Association, Chicago, IL.

Levitt, S. D. (1997). Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review, 87,* 270–290.

Experiments, Laboratory

Laboratory experiments are used in political science to empirically test the results and assumptions of formal models. These types of experiments are conducted using human subjects in a specific location, usually in a computer laboratory, where the experimenter has complete control over the experimental environment. These types of experiments often use a computer interface to deliver the experimental stimuli to subjects. For example, laboratory experiments are used to empirically test formal models of elections, committees, bargaining, and other political topics. This entry presents some of the special features and advantages of this method.

Formal models are mathematically defined and derived by assumptions where the assumptions (sometimes) result in an equilibrium prediction. In formal models, the assumptions are of two types: assumptions about institutional and other exogenous factors such as voting rules, bargaining procedures, preference distributions, and so forth, and assumptions about the behavioral choices of the political actors in the context of these exogenous factors. The model makes assumptions about both electoral institutions and voters' rationality, and the predictions are behavioral.

Laboratory experiments are ideal to test formal models since the assumptions of the model can be operationalized in the laboratory, and human subjects can be used to play the role of actors within the model. Hence, researchers can empirically determine how the assumptions and predictions of the model hold up when human subjects play the game. In laboratory experiments, researchers are able to establish controls over the experimental environment that allow them to establish and measure causality, which is how one variable(s) affects another variable(s). In experiments, causality is established by specifying different control variables and then having subjects randomly assigned to various treatments. For instance, if a researcher is interested in how the cost of voting affects voting behavior, the researcher can specify two treatments—one with a voting cost and another without a voting cost. Then, holding all other parameters of the experiment constant, it is possible to compare the two treatments and isolate the effects of voting cost, given the parameters introduced in the experiment.

Most laboratory experiments use a within-subject design as opposed to a between-subject design. In a within-subject design, subjects in the experiment experience all treatments; in a between-subject design, subjects experience only one treatment. In a within-subject design, subjects are randomly assigned to all the treatments over multiple periods. One advantage of a within-subject design is that it greatly increases the data that are collected. Another advantage as opposed to a between-subject design is that it reduces the error variance associated with differences in subject behavior (i.e., in a between-subject design, subjects may not be randomly mixed and might have shared factors that are unknown to the experimenter). Third, a within-subject design reduces repeated game effects, which means that most models are a one-shot game, but experiments usually test the model over multiple periods. To keep the one-shot nature of the model intact, subjects are assigned to different treatments over multiple periods, so that they will experience a different experimental environment in each period. Hence, each period will be a new experience, and subjects should not feel like they are playing the same experiment in each period.

As mentioned above, causality is established by experimental controls and randomization of subjects to different treatments, allowing the researcher to directly observe the impact of one variable on another without the use of sophisticated statistical methods. One important control is the motivation of subjects. Laboratory experiments follow the dictates of induced value theory. This means that subjects in the experiment have monetary incentives, and they are paid earnings based on their performance during the experiment. With this payment structure then, one subject might take an action that earns him $1, while another subject might take a different action that earns her $2. Induced value theory specifies that as long as subjects are sufficiently rewarded monetarily, subject choices in the experiment should be natural.

To enhance a subject's monetary incentive and to control factors that are not controlled for in the experiment (e.g., psychological factors), laboratory experiments use a sterile environment as opposed to a contextual environment. For instance, in an election experiment, instead of having voters vote for two candidates labeled the left and the right candidate, candidates are given generic labels, such as A or B, or Green and Red. The reason is that a subject might vote for a right candidate because the subject is conservative and thinks the right candidate is more aligned with his or her preferences. Hence, the subject might vote for a candidate based on ideological preferences and mistakenly not vote for a candidate based on the monetary structure induced within the experiment.

Using monetary incentives, the experimenter is able to establish controls over the preferences of subjects. Why do researchers need to control the preferences of subjects? Because this allows a subject's behavior to be measured. Generally, subjects are assigned a utility function that maps monetary amounts to various outcomes. For instance, assume that there are three alternatives that a subject confronts during the experiment—A, B, C—where for a subject, A is worth $2, B is worth $1, and C is worth $0. Hence, a subject's preferences over the alternatives are $A > B > C$. In the experiment, a subject's motivation is to make choices that yield the highest possible payoff for the subject. Subjects in the experiment can have conflicting interests, such that another subject might have the preferences $C > B > A$.

By controlling preferences, it is possible to examine a wide range of topics, but let us consider one: strategic and coordinated voting in three-candidate elections. To illustrate, suppose a group of four subjects (Group 1) has the preferences $A > B > C$, another group of four subjects (Group 2) has the preferences $B > A > C$, and a third group of six subjects (Group 3) prefers C and is indifferent between A and B (i.e., $A = B$). If the experimental election assumed plurality voting and all subjects voted for their most preferred alternative, then C would win the election. But note that this is the least preferred outcome of Groups 1 and 2. Hence, members of Groups 1 and 2 have an incentive to vote strategically—that is to vote for a least preferred outcome to get a better one. To accomplish this, either members of Group 1 would have to strategically vote for B or members of Group 2 would have to strategically vote for A. Hence, it is a coordination problem. In experiments Robert Forsythe and colleagues found that without any type of coordination device, voters were unable to coordinate.

Within this framework, it is possible to vary other control variables to determine the impact that other variables have on the ability of voters to coordinate their vote when strategic voting will increase their payoffs. One variable is the type of election rule subjects vote under. Experiments that varied election rules, such as the Borda rule and approval voting, found that as compared with plurality voting, the Borda rule and approval voting allowed subjects to coordinate more easily.

Another important control variable that is established in the laboratory is the information that subjects have about the experimental environment. For instance, in election experiments, it is possible to vary the information that subjects have about candidates (i.e., a voter may know that A is her most preferred candidate but does not know if she prefers B or C). Controlling subject information has been used to experimentally study the classic Downsian model of elections. In this model, political candidates adopt policy positions over a single-issue dimension (i.e., a single line that ranges from 0 to 100). Voters in this setup have single-peaked preferences over the dimension (single-peaked preferences means that voters have an ideal point over the dimension or a best location). Assuming an odd number of voters who know where the candidates are located on the dimension and of candidates who know where voters are located on the dimension, candidates' optimal location or equilibrium is at the position of the median voter—which is the well-known median voter theorem.

Experiments were conducted that relaxed the full-information condition of the model to determine if the median voter result would still hold. Richard McKelvey and Peter Ordeshook limited the information that voters and candidate subjects had about the election environment. They found that when voters only knew the relative location of candidates and candidates had no knowledge of the voters' distribution of preferences, then outcomes generally conformed to the median voter's ideal point.

In addition to the controls over the experimental environment, laboratory experiments allow researchers to study political phenomena that would be difficult to study using real-world data. For instance, Marco Battaglini, Rebecca Morton, and Thomas Palfrey examined a "swing voter's curse" that relates to the paradox of why people vote, since the probability of a single vote affecting the outcome is very small. If there is a cost to voting, then there is no rational reason to vote, but what if a voter pays the cost of voting and goes to the polling station but then does not vote for certain items on the ballot? The swing voter's curse happens when a voter forgoes voting for an item on a ballot and delegates this decision to other voters whom he thinks are more informed. The results showed that uninformed voters do in fact delegate their vote or abstain most of the time in these cases. Hence, knowledge of why voters behave in this manner would be difficult to discern without using laboratory experiments.

One concern that has been expressed about laboratory experiments among nonexperimental researchers is related to validity issues. The reason that experimental researchers face this question is that nonexperimental researchers feel that a computer laboratory is an artificial environment and does not replicate the real world. However, the artificiality of the environment is a desirable property, and it allows the researcher to make precise measurements of behavior within the laboratory. The results in the laboratory can be compared with other studies using different types of data to further enhance validity.

In sum, laboratory experiments are used to test the assumptions and predictions of formal models. Using this methodology, it is easy to measure causality by specifying different control variables and randomly assigning subjects to different treatments. Control over motivation in the experiment allows for the control of preferences and allows for the measurement of behavior. Also, using this method, a researcher is able to examine the properties of political mechanisms that have not been instituted or used very little in the real world.

Finally, laboratory experiments allow the researcher to empirically examine all types of political mechanisms and collect relevant data, which are difficult to obtain using nonexperimental methods.

Rebecca B. Morton
New York University
New York City, New York, United States

Kenneth C. Williams
Michigan State University
East Lansing, Michigan, United States

See also Causality; Comparative Methods; Experiments, Field; Experiments, Natural

Further Readings

Battaglini, M., Morton, R. B., & Palfrey, T. R. (2010). The swing voter's curse in the laboratory. *Review of Economic Studies, 71*(1), 61–89.

Forsythe, R., Myerson, R. B., Rietz, T., & Weber, R. (1996). An experimental study of voting rules and polls in a three-way election. *International Journal of Game Theory, 25,* 355–383.

Kinder, D., & Palfrey, T. (1993). *Experimental foundations of political science.* Ann Arbor: University of Michigan Press.

McKelvey, R. D., & Ordeshook, P. C. (1985). Elections with limited information: A fulfilled expectations model using contemporaneous poll and endorsement data as informational sources. *Journal of Economic Theory, 36,* 55–85.

Morton, R. B., & Williams, K. C. (2008). Experimentation in political science. In J. Box-Steffensmeier, D. Collier, & H. Brady (Eds.), *The Oxford handbook of political methodology* (pp. 339–356). Oxford, UK: Oxford University Press.

Morton, R. B., & Williams, K. C. (2010). *Experimental political science and the study of causality: From nature to the lab*. Cambridge, UK: Cambridge University Press.

Experiments, Natural

The importance of natural experiments—analyses of real-world situations that attempt to employ some type of random assignment—is in their contribution to addressing confounding, a pervasive problem in the social sciences. Consider the obstacles to addressing the following hypothesis: Extending property titles to poor land squatters boosts access to credit markets and promotes beliefs in individual political efficacy, thereby fostering socioeconomic development. To test this idea, researchers might compare poor squatters who possess land titles with those who do not. Yet confounding may be a problem because differences in individual attitudes and behaviors could in part be due to factors—such as family background—that also make certain poor squatters more likely to acquire titles to their property. This entry examines both the strengths and potential limitations of natural experiments and how they fit among the various research strategies available to political scientists.

It is useful to contrast natural experiments with both conventional observational studies and true experiments. In the first approach, investigators seek to control for potential confounders in observational (nonexperimental) data. For instance, they may compare titled and untitled squatters within strata defined by measures of family background. At the core of conventional quantitative methods is the hope that such confounders can be identified, measured, and controlled. Yet this is not easy to do. Even within the strata defined by family background and intelligence, there may be other difficult-to-measure confounders—say, determination—that are associated with obtaining titles and that also influence economic and political behaviors.

A second approach would be to use a randomized controlled experiment to estimate the effects of land titling. Subjects could be randomly assigned to receive titles or not; family background, determination, and other possible confounders would then be equivalent, on average and up to random error, across these two groups. Thus, large posttitling differences would then be credible evidence for a causal effect of land titles. However, experimental research in such contexts may be expensive, impractical, or unethical.

Scholars therefore increasingly make use of natural experiments—attempting to identify and analyze real-world situations in which some process of random or *as-if* random assignment places cases in alternative categories of the key independent variable. In the social sciences, this approach has been used to study the relationship between lottery winnings and political attitudes, the effect of voting costs on turnout, the impact of quotas for women village councilors on public goods provision in India, and many other topics. In the health sciences, a paradigmatic example comes from John Snow's 19th-century tests of the hypothesis that cholera is waterborne.

Natural experiments share one crucial attribute with true experiments and partially share a second attribute. First, outcomes are typically compared across subjects exposed to a treatment and those exposed to a control condition (or a different treatment). Second, in partial contrast with true experiments, subjects are often assigned to the treatment not at random but rather as-if at random (though sometimes true randomization occurs, as in lottery studies). Given that the data come from naturally occurring phenomena that often entail social and political processes, the manipulation of the treatment is not under the control of the analyst; thus, the study is observational.

However, a researcher carrying out this type of study can often make a credible claim that the assignment of nonexperimental subjects to treatment and control conditions is as good as random. This distinguishes natural experiments from "quasi experiments," in which comparisons are also made across treatment and control groups, but nonrandom assignment to treatment is a key feature of the designs.

Yet how can the claim of as-if random assignment in natural experiments be validated? And how much leverage do natural experiments in fact provide for causal inference? These questions are discussed below, after the initial example on land titling is discussed at greater length.

How Do Property Rights Affect the Poor?

An interesting social-scientific example comes from a study of how land titles influence the socioeconomic development of poor communities. In 1981, urban squatters organized by the Catholic Church in Argentina occupied open land in the province of Buenos Aires, dividing the land into parcels that were allocated to individual families. A 1984 law, adopted after the return to democracy in 1983, expropriated this land with the intention of transferring titles to the squatters. However, some of the original landowners challenged the expropriation in court, leading to long delays in the transfer of titles to some of the squatters. By contrast, for other squatters, titles were granted immediately.

The legal action therefore created a (treatment) group of squatters to whom titles were granted promptly and a (control) group to whom titles were not granted. The authors of the study found subsequent differences across the two groups in standard social development indicators: average housing investment, household structure, and educational attainment of children. In contrast, the authors do not find a difference in access to credit markets, which contradicts a well-known theory that the poor will use titled property to collateralize debt. They also found a positive effect of property rights on self-perceptions of individual efficacy. For instance, squatters who were granted land titles—for reasons over which they apparently had no control—disproportionately agreed with statements that people get ahead in life due to hard work.

Is this a valid natural experiment? The key claim is that land titles were assigned to the squatters as-if at random, and the authors present various kinds of evidence to support this assertion. In 1981, for example, the eventual expropriation of land by the state and the transfer of titles to squatters could not have been predicted. Moreover, there was little basis for successful prediction by squatters or the Catholic Church organizers of which particular parcels would eventually have their titles transferred in 1984. Titled and untitled parcels sat side-by-side in the occupied area, and the parcels had similar characteristics, such as distance from polluted creeks. The authors also show that the squatters' characteristics, such as age and sex, were statistically unrelated to whether they received titles—as should be the case if titles were assigned at random. Finally, the government offered equivalent compensation—based on the size of the lot—to the original owners in both groups, suggesting that the value of the parcels does not explain which owners challenged expropriation and which did not. On the basis of extensive interviews and other qualitative fieldwork, the authors argue convincingly that idiosyncratic factors explain some owners' decisions to challenge expropriation and that these factors were unrelated to the characteristics of squatters or their parcels.

The authors thus present compelling evidence for the equivalence of treated and untreated units. Along with qualitative evidence on the process by which the squatting took place, this evidence helps bolster the assertion that assignment is as-if random. Of course, assignment was not randomized, so the possibility of unobserved confounders cannot be entirely ruled out. Yet the argument for as-good-as-random assignment appears compelling. Note that the natural experiment plays a crucial role. Without it, the intriguing findings about the self-reinforcing (not to mention self-deluding) beliefs of the squatters could have been explained as a result of unobserved characteristics of those squatters who did or did not successfully gain titles.

A Framework for Evaluating Natural Experiments

How much leverage for causal inference do natural experiments in fact provide? To address this question, it is helpful to discuss three dimensions along which natural experiments may vary:

1. plausibility of an as-if random assignment;
2. credibility of the statistical models, which is closely connected with the simplicity and transparency of the data analysis; and
3. substantive relevance of the intervention—that is, whether and in what ways the specific contrast between treatment and control provides insight into a wider range of important issues and contexts.

First, the key claim—and the definitional criterion—for a natural experiment is that treatment assignment is as good as random. Yet one finds marked variation in the plausibility of this claim among studies that claim to use natural experiments.

How can the assertion of as-if random be at least partially validated? First, it should be supported by the available empirical evidence—for example, by showing equivalence on relevant pre-treatment variables (those whose values were determined before the intervention took place) across treatment and control groups, as would occur on average with true randomization. Qualitative knowledge about the process by which treatment assignment takes place can also play a key role in validating a natural experiment. The authors of the studies on land titling use both quantitative comparisons and qualitative evidence about the process of organizing squatters' settlements to validate the claim of an as-if random assignment.

A second dimension along which natural experiments may vary is in the credibility of the statistical models used to analyze the data. As with true experiments, as-if random assignment implies that both known and unknown confounders are balanced (in expectation) across treatment and control groups, obviating the need to measure and control for confounding variables. This has the great advantage of permitting the use of simple analytic tools—for example, comparisons of means or percentages across the treatment and control groups—to make causal inferences. In principle, the simplicity and transparency of the statistical analysis provides natural experiments with an important advantage, relative to the conventional quantitative methods that have in recent years incurred substantial criticism from leading methodologists.

In practice, greater credibility of statistical models is not inherent in all studies that claim to use natural experiments. One recent survey of leading examples from political science and economics found that about half of the studies failed to present simple, unadjusted difference-of-means tests (in addition to any auxiliary analyses). Of course, in less-than-perfect natural experiments, in which the plausibility of as-if random is perhaps impeachable, researchers may feel compelled to control for potential confounders they can measure. Yet any substantial changes after adjustment likely point to a lack of as-if random assignment—because randomization would ensure that control variables are independent of treatment assignment. Post hoc statistical fixes can also lead to data mining, with only "significant" estimates of causal effects making their way into published reports. (Researchers also sometimes use multivariate regression to reduce the variability of treatment effect estimators; yet variance may be higher or lower after adjustment, and the standard errors calculated using the usual regression formulas do not apply.) Thus, to bolster the credibility of the statistical models used in natural-experimental designs, analysts should report unadjusted difference-of-means tests, in addition to any auxiliary analyses.

A third dimension along which natural experiments can be classified is the substantive relevance of the intervention. Here we ask, "To what extent does as-if random assignment shed light on the wider social-scientific, substantive, and/or policy issues that motivate the study?" Answers to this question might be a cause for concern, for a number of reasons. For instance, the type of subjects or units exposed to the intervention might be more or less like the populations in which we are most interested. In lottery studies of electoral behavior, for example, levels of lottery winnings may be randomly assigned among lottery players, but we might doubt whether lottery players are like other populations (say, all voters). Next, the particular treatment might have idiosyncratic effects that are distinct from the effects of greatest interest. To continue the same example, levels of lottery winnings may or may not have similar effects on, say, political attitudes as income earned through work. Finally, natural experimental interventions (like the interventions in some true experiments) may "bundle" many distinct treatments or components of treatments, which may limit the extent to which this approach isolates the effect of the explanatory variable about which the researcher cares most. Such ideas are often discussed under the rubric of "external validity," but the issue of substantive relevance involves a broader question—that is, whether the intervention in fact yields causal inferences about the real causal hypothesis of concern.

To tie this discussion together, we can imagine a cube, in which the three axes are defined by these three dimensions:

1. plausibility of as-if random assignment,
2. credibility of statistical models, and
3. substantive relevance of intervention.

In the front lower-left corner of the cube, we find those natural experiments that offer the least plausibility, credibility, and substantive relevance. In the back upper-right corner of the cube, we find those studies with the most plausibility, credibility, and relevance. The process of achieving a strong research design—in which natural experiments provide substantial leverage for causal inference—may be understood as the process of moving from the front lower-left, "weak research design" corner of the cube, to the back upper-right, "strong research design" corner of the cube.

Two points are important to make in closing. First, there may be trade-offs in seeking to design a strong natural experiment—that is, in moving to the back upper-right corner of the cube. Different studies may manage the trade-off among these three dimensions in different ways, and which trade-offs are acceptable (or unavoidable) may depend on the question being asked.

Second, deep substantive knowledge, and a combination of quantitative and qualitative analyses, can help analysts better achieve success along all the three dimensions of the framework described above. Consider the studies of squatters in Argentina. There, substantive knowledge was necessary to recognize the potential to use a natural experiment to study the effect of land titling, and many field interviews were required to probe the plausibility of as-if randomness—that is, to validate the research design. Fieldwork can also enrich analysts' understanding and interpretation of the causal effects they estimate.

In sum, many modes of inquiry may contribute to successful causal inference using natural experiments; ultimately, the right mix of methods substantially depends on the research question involved. Natural experiments, like regression analysis, do not provide a technical quick fix to the challenges of causal inference. In every study, analysts are challenged to think critically about the match between the assumptions of models and the empirical reality they are studying. Natural experiments are valuable to the extent that they build on real substantive knowledge and appropriate methodological craftsmanship and a full awareness of the trade-offs inherent in this style of research.

Thad Dunning
Yale University
New Haven, Connecticut, United States

See also Causality; Experiments, Field; Experiments, Laboratory; Mixed Methods; Quantitative Versus Qualitative Methods

Further Readings

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Boston: Houghton Mifflin.

Di Tella, R., Galiani, S., & Schargrodsky, E. (2007). The formation of beliefs: Evidence from the allocation of land titles to squatters. *Quarterly Journal of Economics, 122,* 209–241.

Dunning, T. (2010). Design-based inference: Beyond the pitfalls of regression analysis? In H. E. Brady & D. Collier (Eds.), *Rethinking social inquiry: Diverse tools, shared standards* (2nd ed., pp. 273–311). Lanham, MD: Rowman & Littlefield.

Freedman, D. A. (2010). Statistical models and shoe leather. In D. Collier, J. S. Sekhon, & P. B. Stark (Eds.), *Statistical models and causal inference* (pp. 45–62). New York: Cambridge University Press. (Original work published 1991)

Galiani, S., & Schargrodsky, E. (2004). The health effects of land titling. *Economics & Human Biology, 2,* 353–372.

Snow, J. (1936). *Snow on cholera.* London: Oxford University Press. (Original work published 1855)

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *Journal of Educational Psychology, 51*(6), 309–317.

F

Factor Analysis

Factor analysis is a well-established method that attempts to measure latent constructs such as attitudes or values, which cannot be observed directly. The standard model assumes that the measured variables (items or indicators) are linear additive functions of the unobserved (latent) factors and the error component. This is called the common-factor model. The general system of equations for this model is

$$y_j = \lambda_{j1}\eta_1 + \lambda_{j2}\eta_2 + \cdots + \lambda_{jm}\eta_m + \varepsilon_j,$$

where y_j represents the $1, \ldots, j$ observed variables measured on a sample of n independent subjects; $\eta_1, \ldots, \eta_m$ represent the $1, \ldots, m$ latent constructs (factors) in the model; $\lambda_{j1}, \ldots, \lambda_{jm}$ represent the factor loading (partialized regression coefficient) relating variables $1, \ldots, j$ to the first to mth factors; and $\varepsilon_1, \ldots, \varepsilon_j$ stands for the error component (uniqueness). It is assumed that the error component of one indicator or item is independent of all factors and of all error components of the other items. In this entry, the two major forms—exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)—are presented.

An example with five items ($y_1, \ldots, y_5$), two factors (η_1, η_2), and five random measurement errors ($\varepsilon_1, \ldots, \varepsilon_5$) is given in Figure 1; ϕ_{12} is the symbol for the covariance between the Factors 1 and 2.

The variance of the unique component can be further decomposed as $\varepsilon = \varepsilon_s + \varepsilon_e$, where ε_s is the specific error variance due to the particular item (e.g., a specific item wording) and ε_e the random measurement error. In most cross-sectional studies, one cannot distinguish between these two kinds of errors because they are not separately identified. It is possible to separate these components only by using special designs like multitrait–multimethod or panel studies.

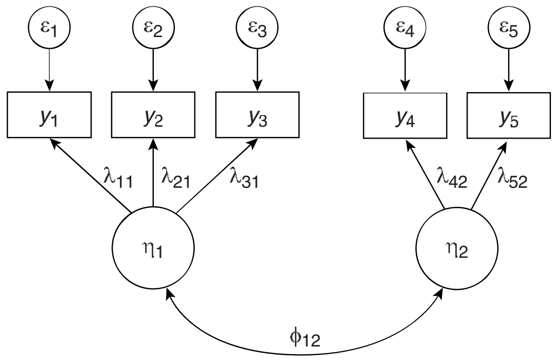


Figure 1 Visualization of a Factor Model With Five Observed and Two Latent Variables

Generally, one has to differentiate between EFA and CFA. EFA is used when there is no or not sufficient a priori knowledge about the number of factors and the relationship between items and constructs. CFA, in contrast, is used when researchers have concrete assumptions about the measurement model.

Exploratory Factor Analysis

EFA is a method used to detect the optimal number of factors that accounts for the correlation of items. Each factor is interpreted and named based on the items that have high loadings on this factor. Its character is inductive as the number of factors, the amount of correlations between factors, and the assignments between items and factors are performed empirically without a precise deductive theoretical model. There is no unique solution for the relation between observed items and latent factors. There are two different approaches to the factor model—(a) common-factor model and (b) principal component analysis—and each implies different substantive assumptions. In the common-factor model, it is assumed that every item is measured with some error. Therefore, the common-factor model is most appropriate in survey research, where items are always measured with some error. However, in the principal component model, the researcher assumes that there is no measurement error in the items. All analyses are based on correlation matrices as input.

EFA is performed in five steps:

1. *Factor extraction:* The method that is most often used is the method of principal axis factoring and is based on the computation of the eigenvalues and vectors of the correlation matrix between all measured variables. Its goal is the maximization of the variance of each successively extracted factor.

2. *Number of factors:* Different procedures determine the number of factors: (a) According to the Kaiser-Guttman criterion, the number of factors should correspond to the number of eigenvalues of the full input correlation matrix that exceed one. (b) The scree plot is a graph of each eigenvalue plotted in descending order. The number of factors is determined by visual inspection by observing whether, suddenly, the plot shows no real difference between the last two eigenvalues in the scree plot. (c) In parallel analysis, a second factor analysis is calculated with a random data set with the same numbers of variables and cases as in the original analysis. Only the factors with higher eigenvalues than the factors of the random data are accepted. (d) In maximum likelihood estimation, one studies the amount of residuals left over after introducing more factors and evaluates them by goodness-of-fit tests of the whole model.

3. *Communality estimation as a measure for the common variance of every item:* The most often used method for an initial estimate is the squared multiple correlation of one item with all other items. Depending on the number of factors extracted, the final solution contains the explained variance of every item.

4. *Factor rotation:* The goal of rotation is to achieve "a simple structure" that allows a good interpretation of the resulting solution. Rotations can be orthogonal, assuming that there are no correlations between factors (e.g., varimax, quartimax, and equimax), or oblique (e.g., promax, oblimin), assuming that there exist correlations between factors. In the case of correlated factors, the factor loadings of items on factors contained in the pattern factor matrix represent standardized partialized regression coefficients. In the case of noncorrelated factors, these coefficients are simply correlation coefficients between items and factors. The factor structure matrix contains the simple bivariate correlations between items and constructs. The values differ from the factor pattern matrix only if factors are correlated.

5. *Factor score estimation:* As a final step, one may be interested in estimating the factor scores. The factor scores are often used as weights for raw scores in the computation of indices that are used in subsequent analyses.

Confirmatory Factor Analysis

In contrast to EFA, in CFA, a theoretical model is needed that contains a series of a priori hypotheses. These hypotheses should be drawn from the substantive literature. For example, in the theory of values of Shalom Schwartz, 10 values are postulated and theoretically derived. Therefore, the 10 theoretical constructs can be formalized and tested as latent variables in a CFA model. However, it is not definitely determined in the literature which and how many dimensions and items are adequate to measure the concepts of nationalism and patriotism and whether one can differentiate between them. The hypotheses in a CFA model refer to the following aspects of a measurement model: (a) determination of the number of factors, (b) assumption regarding whether the factors are correlated or not, (c) determination of which items load on which factor and where to set loadings a priori to zero, and (d) assumptions about the correlations of the errors. CFA allows the specification and testing of these assumptions. It is possible to fix parameters to a certain value or to constrain parameters that are set equal to other parameters. One assumes that the expected value of random measurement error is zero ($\varepsilon_s = 0$) and the correlation of the random measurement error with the latent variable (factor) is also zero. Furthermore, it is necessary to fix one of the loadings to one or to standardize the variance of the latent variable to reach identification of the model, that is, to achieve a unique solution.

To illustrate this, in this entry, we use data from the International Social Survey Programme (ISSP) National Identity module collected in 2003, which contains items to measure nationalism and constructive patriotism. Theoretically, it has been discussed whether there are two different latent variables (factors)—namely, nationalism and constructive patriotism. Nationalism was measured on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The questions were worded as follows: (1) "The world would be a better place if people from other countries were more like the [country nationality of the respondent]." (2) "Generally speaking, [the respondent's country] is a better country than most other countries." A high score indicates a high degree of nationalism. Constructive patriotism was measured using the following three items: "How proud are you of [respondent's country] in each of the following:" (3) "The way democracy works." (4) "Its social security system." (5) "Its fair and equal treatment of all groups in society." Responses were given on a 4-point scale ranging from 1 (*not proud at all*) to 4 (*very proud*). The Pearson correlation matrix and standard deviations are provided for the United States in Table 1. Pairwise deletion of missing values was used because the number of missing values was low (less than 5%).

The standardized and unstandardized coefficients were computed using the maximum likelihood estimation procedure in AMOS version 18 software with graphical input (see Figure 2).

Figure 2 shows that the two latent variables (factors)—constructive patriotism and nationalism—are positively correlated ($r = .48$). The standardized factor loadings are all equal to or higher than .50, which is an expression of the sufficient formal validity of the items in the model. Furthermore, "demo," "security," and "fair" are only explained by patriotism, whereas "like us" and "better" are only explained by nationalism. There are no cross-loadings. Therefore, both convergent validity by sufficient factor loadings of the items on their respective factors and divergent validity by having no cross-loadings and a correlation between the

Table 1 Means, Standard Deviations, and Correlations Between Items (U.S. Sample)

		Like Us	*Better*	*Demo*	*Security*	*Fair*
Mean		2.74	1.87	1.80	2.44	2.06
Standard deviation		1.11	0.90	0.68	0.85	0.81
Like us	*r*	1				
	N	1,197				
Better	*r*	.48	1			
	N	1,187	1,197			
Demo	*r*	.16	.23	1		
	N	1,138	1,140	1,150		
Security	*r*	.14	.15	.31	1	
	N	1,156	1,154	1,118	1,167	
Fair	*r*	.23	.21	.35	.31	1
	N	1,153	1,151	1,116	1,129	1,162

Source: Data from International Social Survey Programme (ISSP), National Identity II (2003). Distributor: GESIS Cologne Germany, ZA No. 3910.

Note: r = Pearson's correlation coefficient; correlations between indicators of the same construct are shaded.

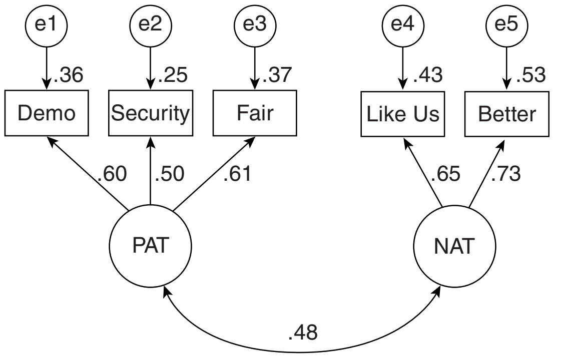


Figure 2 Results From a Confirmatory Factor Analysis for the U.S. Sample (Standardized Solution)

Source: Data from International Social Survey Programme (ISSP), National Identity II (2003). Distributor: GESIS Cologne Germany, ZA No. 3910.

Note: PAT = patriotism; NAT = nationalism.

concepts that is not too high have been established. The explained variances of the observed variables (given at the edge of the rectangles) are between .25 for "security" and .53 for "better." The unstandardized factor loadings and their standard errors are reported in Table 2. The model fit (not reported) is satisfactory according to standard criteria.

One important generalization of CFA is multiple group confirmatory factor analysis (MGCFA), which allows us to test the invariance of factor loadings, factor covariances and variances, measurement errors, intercepts of observed variables, and latent means. This procedure is especially relevant for comparative social research and politics—for example, in cross-national comparisons to establish measurement equivalence across countries. The invariance of the factor model presented above for nationalism and constructive patriotism was tested in 34 countries. According to the global fit measures, the factor loadings were not substantially different between the countries in the ISSP 2003. Therefore, one can conclude that the necessary condition for the equality of meaning of the constructs in the 34 countries is given, and the relations can thus be compared on the theoretical level.

Furthermore, in both EFA and CFA, higher order factors can be estimated. This means that the correlation between factors is explained with one or several higher order factors. Further topics are nonlinear CFA, CFA with categorical and nonnormal indicators, the relationship to item–response theory, application of Bayesian estimation, CFA with longitudinal data, multilevel CFA, and mixture CFA.

Table 2 Unstandardized Factor Loadings and Standard Errors (U.S. Sample)

		PAT	*NAT*
Like us	λ	0	1.000
	SE	0	[a]
Better	λ	0	1.106
	SE	0	0.130
Demo	λ	0.822	0
	SE	0.078	0
Security	λ	0.873	0
	SE	0.085	0
Fair	λ	1.000	0
	SE	[a]	0

Source: Data from International Social Survey Programme (ISSP), National Identity II (2003). Distributor: GESIS Cologne Germany, ZA No. 3910.

Notes: λ = unstandardized factor loadings; PAT = patriotism; NAT = nationalism; *SE* = standard error of λ.

a. No *SE* is estimated because 1.0 is a fixed parameter, which is necessary for identification of the model.

Peter Schmidt
National Research University
Higher School of Economics (HSE)
Moscow, Russia

Johannes Herrmann
Justus-Liebig-Universität Gießen
Gießen, Germany

See also Analysis of Variance; Statistics: Overview

Further Readings

Bandalos, D. L., & Boehm-Kaufman, M. R. (2009). Four common misconceptions in exploratory factor analysis. In C. Lance & R. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 61–88). New York: Routledge.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.

Davidov, E. (2009). Measurement equivalence of nationalism and constructive patriotism in the ISSP: 34 countries in a comparative perspective. *Political Analysis, 17*(1), 64–82.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*(2), 183–202.

Mulaik, S. A. (2009). *Foundations of factor analysis*. London: Chapman & Hall.

Fair Division

Fair division analyzes how to divide divisible or indivisible goods among two or more players in order to satisfy certain criteria of fairness. The problem of fair division goes back at least to the Bible. Abraham and Lot had to decide who would get Canaan and who Jordan; Solomon had to decide which of two women was the mother of a disputed baby. What was a fair solution in each case? This entry first discusses the criteria of fair division, followed by specific procedures and some applications in political science and everyday life.

The oldest known procedure for dividing a single divisible good, such as a cake or land, between two players is "I cut, you choose," or divide-and-choose. The same procedure can be used if there are multiple indivisible goods: The divider partitions the items into two piles, and the chooser selects one pile. If the divider has no knowledge of the chooser's preferences, the divider should divide the items 50/50. That way, whichever pile the chooser selects, the divider is assured of getting 50%. The chooser, by contrast, will get more than 50% if she thinks that the two piles are unequal and selects the one that she thinks is more valuable. If, however, the divider knows the chooser's preferences, he can exploit this information to make one pile slightly more valuable than 50% for the chooser, so the chooser will select that, keeping for the divider the pile he values more (assuming that he values items different from what the chooser values). It is important to note here that fairness does not refer to an "objective" criterion of equality but agreement according to one's subjective preferences.

Criteria of Fair Division

Is divide-and-choose fair? To make an assessment, consider the following criteria for determining what a fair share is:

Proportionality: If there are n players and they are each equally entitled to the items, a division is proportional if each player thinks that he or she received at least $1/n$ of the total value. Proportionality can be traced back to Aristotle, who argued in his book *Ethics* that goods should be divided in proportion to each player's contribution.

Envy Freeness: If no player is willing to give up his portion in exchange for the portion another player receives, this player will not envy any other player. In two-player disputes, a division is envy free if and only if it is proportional. To see why this is so, assume that a settlement is proportional, so you think you are getting at least 1/2 of the total value of all the items. Will you envy me? Not if you think that you have at least 1/2, because then you must think that I have at most 1/2. Symmetrically, if I think I have at least 1/2, then I will not envy you, so the settlement is envy free. Conversely, if the settlement is envy free, then both of us think that each is getting at least 1/2; otherwise, at least one of us will envy the other for getting more than 1/2. Thus, if there are only two players, proportionality and envy freeness are equivalent.

In the case of three players, however, envy freeness is more demanding than proportionality. For example, I may think I'm getting 1/3, but if I think you're getting 1/2 (e.g., because the third player, in my eyes, is getting only 1/6), then I will envy you. By contrast, if an allocation among three players is envy free, then I must think I received at least 1/3 (otherwise, I would think the others together received more than 2/3, and I would envy the one or both players who received more than 1/3). Hence, an envy-free allocation is always proportional, even if there are more than two players, but a proportional allocation is not necessarily envy free.

Equitability: A division is equitable if all players think that they receive exactly the same value above $1/n$. Equitability is an aspect of satisfaction more subtle than envy freeness. To illustrate it in a divorce settlement, suppose you think you got 51%

of the marital property but your spouse thinks she got 90% (because she had little interest in what you got). While you do not envy the portion of your spouse—in your eyes, she got only 49%—you may envy her happiness: She got more of what she wanted than you did of what you wanted.

Equitability means that both players think that they receive the same fraction of the total as each of them values different items. Coupled with envy freeness, it means not only that both get more than 50% but also that both exceed 50% by the same amount. Thus, if both spouses think that they receive 70% of the total value, they are equally pleased by their allocations. Equitability may be a difficult property to ascertain. How does one measure whether both players are equally happy with their allocations? In fact, it generally requires an interpersonal comparison of utilities.

Efficiency: An allocation is efficient if there is no other allocation that is better for some player without being worse for some other player. Efficiency by itself—that is, when not linked with properties such as proportionality, envy freeness, or equitability—is no guarantee that an allocation will be fair. For example, an allocation that gives everything to me and nothing to you is efficient: Any other allocation will make me worse off when it makes you better off. It is the other properties of fairness, combined with efficiency, that ensure that the maximum possible value is distributed according to the aforementioned properties.

Fair-Division Procedures

The modern mathematical theory of fair division has its roots in the 1940s, particularly in the work of Polish mathematicians. They proposed different procedures for dividing both divisible goods such as a piece of land and indivisible goods such as a car, boat, or house. Procedures that involve dividing divisible goods include a plethora of cake-cutting schemes, some of which use "moving knives" that parties stop when they believe they have a fair share; others involve trimming pieces from a cake, or a set of divisible goods, to construct fair shares. Many of these procedures are quite esoteric and, therefore, impractical to apply. Besides divide-and-choose, probably the two most feasible procedures are the following:

1. *Strict and balanced alternation:* Strict alternation is simply taking turns: You pick an item, then I pick one, you choose again, and so on. Of course, going first can be a huge advantage; giving extra choices to compensate for going second can reduce, if not eliminate, this advantage. A specific way of balancing choices out yields a procedure called balanced alternation. If there are, say, four items to be divided among Players A and B, the order of choice under balanced alternation is ABBA. Balanced alternation can be extended to more than two players and any number of items they may wish to divide.

2. *Adjusted winner:* Two players begin by independently distributing a total of 100 points across all the items to be divided, depending on the relative value they attach to each. Thus, if you consider a certain item to be worth 1/4 of the total value of everything to be divided, then you would put 25 points on it. The term *winner* in "adjusted winner" comes from the next step: Each player is temporarily given the items on which it placed more points than its opponent. Then, the *adjusted* part comes in: Items are transferred from the initial winner to the initial loser in a certain order until the point totals of the two players are equal. This order of transfer, which requires splitting at most one item (it could also be shared or sold and the proceeds divided), guarantees that the final allocation will be envy free, equitable, and efficient if the players are truthful. To be sure, players may try to manipulate a procedure to their advantage by giving false information about their preferences. However, most fair-division procedures provide certain guarantees, even against players that are exploitative.

To return to divide-and-choose, it was shown that the divider can guarantee that he receives a tied-for-largest portion, whatever the chooser selects, by dividing the items 50/50 for himself. Moreover, the chooser can do at least as well by selecting what she considers the larger pile, so this procedure is envy free. However, divide-and-choose may be neither efficient nor equitable. Thus, both players may do better with some other allocation, making the division inefficient. In the case of a chocolate and vanilla cake, for example, if the divider does not know which flavor the chooser favors, the divider will cut it so each pile comprises 50% chocolate and 50% vanilla. But if, in fact, the divider likes chocolate and the chooser likes vanilla, it would be better for both players if the cut were made so as to separate the chocolate from the vanilla, giving each player an opportunity to get the preferred flavor. Divide-and-choose may be inequitable because the chooser could receive a piece she considers larger than 50% when the cutter receives only 50%. Of course, the tables can be turned if the cutter has information about the chooser's preferences and uses it to get a substantially greater than 50% portion for himself when the chooser must settle for only slightly more than 50%.

Applications

The procedures discussed so far offer an algorithmic approach to fair division: A set of rules is specified; subject to the rules, the players select strategies (e.g., the divider cuts the cake into two equal portions) that give each certain guarantees about the resulting allocation (e.g., it will be envy free). Political scientists, sociologists, and applied economists have taken a more empirical tack, seeking to determine the conditions under which fairness and departures from it occur in the world and what consequences they have for people and institutions. Psychologists, especially, have paid heed to how perceptions of fairness impinge on people's attitudes and affect their behavior.

Theoretical economists and game theorists have been more interested in finding axioms that characterize fairness. Their models are often nonconstructive in the sense that they establish the existence—or sometimes the nonexistence—of an allocation satisfying certain properties. (The problem of nonexistence is particularly acute in the case of indivisible goods.) But even models that show existence may provide no clue as to how to construct the desired allocation. Many analysts are interested in fairness because they desire to help people settle their differences amicably. Numerous books have been written that purport to show how to achieve win–win solutions, whereby players win—or at least do not lose—on the issues they consider most important.

Ultimately, players will want to know on which issues they will win, on which they will lose, and on which they will have to compromise. For this purpose, fair-division procedures that can guarantee that the disputants do as well as possible in realizing all the win–win potential that is available can be extremely useful. From a normative perspective, fair-division procedures can help disputants reduce the frustration, anger, and occasional violence that accompany escalating demands and endless haggling. By facilitating the disputants' bringing their own closure to a dispute, fair-division procedures help them avoid the arbitrary imposition of a settlement or continuing impasse. Thereby, these procedures tend to level the playing field, so players with greater resources or capabilities are prevented from gaining unduly large shares. In this sense, fair-division procedures offer a means to prevent exploitation of the weak by the strong and to minimize the effects of power differences.

Steven J. Brams
New York University
New York City, New York, United States

See also Bargaining; Conflict Resolution; Game Theory; Rational Choice

Further Readings

Brams, S. J. (2008). *Mathematics and democracy: Designing better voting and fair-division procedures*. Princeton, NJ: Princeton University Press.

Brams, S. J., & Taylor, A. D. (1996). *Fair division: From cake-cutting to dispute resolution*. New York: Cambridge University Press.

Brams, S. J., & Taylor, A. D. (1999). *The win-win solution: Guaranteeing fair shares to everybody*. New York: W. W. Norton.

Moulin, H. (2003). *Fair division and collective welfare*. Cambridge: MIT Press.

Robertson, J., & Webb, W. (1998). *Cake-cutting algorithms: Be fair if you can*. Natick, MA: A. K. Peters.

FASCISM

Fascism has become a generic term used to represent a political movement that developed between the two world wars principally in Europe, with a few extensions notably in Latin America. This term was employed for the first time as a partisan label in Italy in 1919 by Benito Mussolini, who appealed to younger veterans to establish *Fasci di combattimento*. In Italy, a new type of political organization emerged, developed, and was eventually imposed—one that more or less indirectly foreshadowed and inspired a set of movements and political regimes, including Adolf Hitler's National Socialism in Germany, Léon Degrelle's Rexisme in Belgium, Oswald Mosley's Fascist Union in Great Britain, Corneliu Codreanu's Iron Guard in Romania, Ferenc Szàlasi's Arrow Cross Party in Hungary, José Antonio Primo de Rivera's Falange in Spain, Plinio Salgado's Integralist Action in Brazil, and Jacques Doriot's French Popular Party in France. Within this group, important differences persisted among movements. Some of them remained minority opposition forces, while others formed political regimes; some ruled in coalition, while others gave rise to dictatorships based on a single party. There are also important differences in the ideology or the social basis of these fascist-inspired movements as these emerged and took shape in very different socioeconomic and cultural contexts. Important variables include the degree to which their setting was industrialized and secular and the extent of nationalization and politicization of the masses. It is nevertheless possible to classify a group of movements and regimes into a category labeled *fascism*. These share some major typical traits, formed in the context of the political laboratory represented by the Blackshirts in Italy. This entry analyzes fascism through its connections with World War I, its ideology, and finally the way in which it produced a new kind of regime.

The Decisive Influence of World War I

The first common feature lies in relation to war. In Europe, World War I is indeed the cradle of fascism. It represents a critical element of context, for both the birth of fascist leaders and the organizational and ideological forms that characterized this political movement. Whether in Italy or in Germany, the most committed elements of fascism and National Socialism were, for the most part, young veterans for whom the war proved to be a fundamental initiation experience. Like their respective leaders, Mussolini and Hitler, they supported the war and, above all, gloried in the "spirit of the trenches"; they also suffered from postwar demobilization. Their political trajectory, their values, the ideology they claimed, the political organizations they would frame, and the regimes they sometimes succeeded in creating cannot be understood outside of this initial context: The Great War led to the brutalization of societies.

The influence of the war could already be seen in the type of partisan organizations set up by those who would become fascist leaders. First in Italy, then in Germany and other European countries, they forged a new type of organization: the militia party. This organizational model was a synthesis of the mass party, previously unique to the labor movement, and of the military organization systems developed in the front lines of World War I. The result was a form of political organization unlike any other, driven by agents who, in the case of Italy, came mainly from the middle class, often fought in the ranks of the extreme Left (left-wing Socialist Party, revolutionary syndicalism), and were most often veterans of the Italian Army elite troops (Arditi). This militia party was designed to assemble an avant-garde group organized around a military model: uniforms, strict hierarchical organization, inculcation of discipline, and maintenance of military sociability modeled on the camaraderie of the front. Thanks to this organization, the fascists introduced methods of war into the political arena. In the context of latent civil war that characterized immediate postwar Italy and parts of Europe in the 1920s as well as the climate of political radicalization caused by the crisis of the early 1930s, the fascist militias (including death squads, Storm Troopers, green shirts, etc.) carried out violent attacks on labor movement organizations, which included attacks on the Socialist and Communist parties' public meetings, ransacking of left-wing municipalities, burning down local trade union offices, and intimidating the red league of farmers. During this fierce struggle against labor organizations, the fascists won the sympathy and even the active support of some economic elites and ruling circles, a development that would greatly facilitate their coming to power in a coalition government with Center and Right traditional parties.

Fascist Ideology

These militia parties mobilized their members around a set of values, creeds, and myths that formed the common ideological basis of fascism. This was primarily an ideology of action that found its immediate extension in a style and in a quite singular aesthetic. As in organizational matters, fascisms practiced hybridization in ideological matters. They created a synthesis of influences from the ultra-Right (the most radical nationalist and racist standards) and influences from the ultra-Left (e.g., revolutionary syndicalism inspired by Sorel). France was one of the first laboratories of this new ideological alchemy. The result was a cult of action, violence, and youth and a revolutionary rupture with the parliamentary system and liberal society. Fascist ideology initially involved hostility to the political currents of the Enlightenment; liberalism and Marxism were its targets of choice. It was based on a radical denunciation of social egalitarianism as well as of bourgeois society, presenting itself as a third-force ideology, different from both capitalism and communism. On economic and social development, this third position would find its solution in the corporatist principle of harmonious cooperation between labor and capital for the benefit of the nation; nationalism—not class consciousness—was the backbone of the fascist ideology. This ideology, accordingly, was organized around the myth of the nation conceived as an organic and compact community. Fascism's mission was to purify the nation politically, anthropologically, and even racially with a view to the assertion of its power. Most fascist movements, beginning of course with National Socialism, put racism and anti-Semitism at the center of their doctrine and political plan. Even Italian fascism, which was initially based on a more political conception of the nation, ended in the mid-1930s by putting the issue of race in the foreground and adopting anti-Semitic legislation in 1938. This homogeneous national community was destined to be permanently mobilized for the goal of conquest. The fascist ideology was guided by an imperialist teleology: conquest of Lebensraum in the case of National Socialism and construction of a new world order in the case of Italian fascism. Fascism was not just born from war, it was also born for war: Its intention was to ideologically mobilize the population, with a view to forming a combatant community fit for military conquest.

A Regime of Mass Mobilization

In the Italian and German cases, fascism as a rising power aimed objectively to establish a system of permanent mobilization in which society was seen as an army and the citizen as a soldier. Along with this initiative of a militarized society, the development of fascism into a regime started a political sanctification process; its ideology was introduced into "civil religion," which citizens were called to practice with ardor. This was accompanied by a massive production of liturgies, ceremonies, and political customs. Accordingly, fascism's goal of the "New Man" included both a believer and a soldier: an ideal summed up by one of the most famous programmatic slogans of the Italian fascist regime—"believe, obey, fight."

This mobilizing regime is based on a very specific mechanism. The mass party was the essential element, the central institution of fascist society. Fascist movements came to power in coalition, formally respecting constitutional executives. But very quickly, they implemented strategies to monopolize power by violently destroying labor movement organizations and causing dissolution or taking over the traditional right-wing parties. This process quickly led to a situation of a one-party dictatorship. The dynamics of fascism then operated as a dual process of further state and civil society penetration. The capture of power set in motion a symbiotic process of nationalization of the party and partisan penetration of the state. The state party proved to be an effective tool for the regime's objectives. In particular, it allowed deep penetration into a civil society that was heavily supervised by youth organizations, unions, associations, and governmental agencies in areas such as sports, culture, and recreation. Thanks to the millions of members in fascist organizations, the system could organize the ultrasupervised mobilization of the population, ensure the dissemination of fascist values and ideology, and organize the faith in the regime and its leader. This leader (*Duce*, *Führer*), "Almighty" and "infallible," was also one of the typical elements of fascist domination. His seemingly limitless power was inextricably linked to the strength of the party. This is because the leader had quasi-absolute control over a party that itself had an unrestricted influence over the population. The leader's power was thus a direct reflection of the power of the fascist movement as it became a political regime. Fascist organizations were also formidable instruments that were used to control political conformity. Spread over the whole territory, present in all spheres of social life—in the workplace as well as in recreational activities, in factories as well as in families—they contributed to the establishment of a social and political control exerted on the entire population with intensity unlike any other, finding an equivalent only in Stalin's Soviet Union.

This situation leads to the question of the nature of fascism and, more precisely, its place within the typologies forged by political science to situate the political regimes of the 20th century. The answer to the question of whether fascism should be classified as an authoritarian regime or a totalitarian regime depends, in part, on what principles of logic and restrictive definitions are implemented. If we adopt a restrictive definition of fascism and say that only Italian fascism and Nazism fully qualify, we can consider fascism to be one component of the totalitarian phenomenon. This is indeed a form of domination based on the mobilization of the population through a mass political organization around a civil religion, generating a level of control of individuals that is unprecedented in its intensity, resulting in the eradication of all forms of organized political opposition and functioning even in tandem with any dissenting public expression. Nazism and, to a lesser extent, Italian fascism largely corresponded to this totalitarian type by the time they collapsed under the weight of their military defeats.

Jean-Yves Dormagen
University of Montpellier
Montpellier, France

See also Fascist Movements; Nationalism; Totalitarianism

Further Readings

De Felice, R. (1977). *Interpretations of fascism.* Cambridge, MA: Harvard University Press.

Gentile, E. (2002). *Fascismo: Storia e interpretazione* [Fascism: History and interpretation]. Bari, Italy: Laterza.

Linz, J. (2000). *Totalitarian and authoritarian regimes.* Boulder, CO: Lynne Rienner.

Mosse, G. (1999). *The fascist revolution: Toward a general theory of fascism.* New York: H. Fertig.

Payne, S. (1995). *A history of fascism: 1914–1945.* Madison: University of Wisconsin Press.

Fascist Movements

In the years between World Wars I and II, there were violent fascist movements all over Europe. Nearly every country had at least one; France had six and Switzerland three. The fascist category was quite vague and included, among several competing groups, militant nationalists and veterans of World War I. Their chief antagonists, the socialists and communists, called them fascists—from Benito Mussolini's *fasci di combattimento* (combat squads)—but the criteria for such a label to be used were rather hazy at first. Even after nearly a decade in power, Mussolini himself was reluctant to define the faith of his movement until he allowed Giovanni Gentile, a fascist philosopher, to write his famous article on fascism for the Italian *Encyclopedia Treccani* in 1931. A once-prominent revolutionary socialist leader, Mussolini forged the character of his movement as he went along, seeking violent confrontations with his erstwhile party comrades and accepting the support of whatever group would back him at the time, including former enemies such as big industry, the state, and the Catholic Church. During his struggle for power, he had learned tactics from Vladimir Lenin's violent takeover of the mighty Tsarist Empire. Adolf Hitler in turn sought to imitate Mussolini's methods, and finally, scores of other European fascist movements began to imitate and curry favor with the triumphant Third Reich. Opportunism and constantly changing profiles were typical features of the early fascist movements. Their virulence and relative "purity," however, were quickly watered down and compromised once they participated in governments. Fascist regimes usually sidelined their "old fighters" and left policy to the leadership, the army, and the bureaucracy.

The most important European fascist movements were the Italian National Fascist party (*Partito Nazionale Fascista*, or PNF), the National Socialist (NS) German Workers' Party (*Nationalsozialistische Deutsche Arbeiterpartei*, or NSDAP), the Hungarian Arrow Cross, the Romanian Iron Guard, the Belgian Rexists, and the Spanish JONS/Falange, but there were dozens of others, including many imitators of the most successful fascist movements. Some were more on the side of labor, some with employers. Some claimed to defend the church; most were stridently anti-Semitic and opposed to minorities. Imitation also characterized their use of quasi-military uniforms and ranks, their flags, and their internal organization. The larger movements developed specialized organizations, for example, for youth and women and also for quasi-military enforcement operations and subversion abroad. During World War II, the Third Reich armed forces even established elite Waffen-SS divisions—the SS (originally Hitler's body guard) being otherwise employed for enforcement and concentration camp guard duty. Some movements, such as the NSDAP, also strove to recruit into their ranks corporatist groups, such as the German Labour Front and professional NS associations, for example, a NS Doctors Federation, NS Farmers, and many others to which all German workers, doctors, and farmers, and so on had to belong.

Historical Causes of Fascism

Major historical issues and mid-20th-century conflicts gave rise to most fascist movements. Among the most important was a country's participation in World War I, especially on the losing side. The defeated of the war—the Central Powers (Germany, Austro-Hungary, and Turkey) and others—who felt left out of the distribution of the spoils were fertile ground for organizations of resistance and revenge. It is impossible to imagine the atmosphere surrounding the birth of Italian fascism without mentioning the 600,000 Italian casualties, the sense of having been cheated out of the fruits of victory, and the territorial ambitions of many middle-class Italians. Germans felt similarly stricken by vast casualties and economic losses and obsessed with the loss of German colonies and lands. The war had destroyed four empires (Austro-Hungarian, German, Russian, and Turkish) and led to vast transfers of territory, leaving behind seething resentments and unfulfilled aspirations, sowing hatred and the seed for future conflicts, for example, in the Balkans, and still generating bloodshed and ethnic cleansing today. The so-called axis of fascist powers and movements after World War I was made up of countries and elements violently opposed to the peace settlements of the war, such as the Treaties of Versailles and Trianon. The war also explains the extreme degree of militarization of these fascist movements, the prominent presence of veterans, and the widespread belief among them that quasi-military violence could overpower democratic majority decisions and parliamentary rule. For this purpose, they all sported private armies like the Nazi storm troopers and the Italian Blackshirts and trumpeted their belief in political violence. Fascism thus left a legacy that was most inhospitable to the development of constitutional democracy.

The second reason for the formation of fascist movements was the significant Left–Right schism that had been developing in various European countries for decades. Industrialization, capitalism, and organized labor at different stages of political modernization had reached a point of all-out confrontation by the time of the war. In highly industrialized areas such as Germany and Northern Italy, a vanguard of socialist and communist parties emerged with electoral majorities and was stirred to action by the Russian Revolution of 1917. Tsarist Russia was still mostly agrarian, but the war had brought it to its knees with the onslaught of pacifism, anarchism, and Lenin's socialist forces. The Whites, the military defenders of the monarchy, nobility, and traditional order in the bloody Russian civil war, in many ways resembled the fascist movements of Central and Eastern Europe. In republican Spain, rebellious workers, anarchists, socialists, and communists faced off against capitalists, state forces, the Church, and General Franco's army. In semi-industrial Italy, when the rising socialist challenge culminated in the seizure of factories and agrarian estates, the fascists spearheaded a counterrevolution that was really aimed at establishing their own dictatorship. Everywhere, fascist-like movements were defined by their violent opposition to the inexorable rise of socialist or communist movements, and the brutal use of violence became their trademark.

There were other, later, causes for the fascist eruptions in mid-20th-century Europe: The Great Depression of the 1930s mostly aggravated and deepened the Left–Right schism rather than being the primary cause. Similarly, the rising power of the fascist Axis and its interventions into the domestic politics of small neighboring states—for example, by Mussolini's Italy into Austrian or Hungarian politics—created incentives for fascist movements or profascist realignments. Eventually, fascist puppet regimes came into power in the Balkans and in German-occupied Northern Europe during World War II. Native fascist movements like the Hungarian Arrow Cross, the Romanian Iron Guard, and the Norwegian Nasjonal Samling thus briefly achieved participation in government, although with little popular support.

After World War II, the more prominent interwar fascist movements generated small successor movements such as the Italian Social Movement (*Movimento Sociale Italiano*, or MSI) and the half dozen German neo-Nazi groups. Their members generally were nostalgic "old fighters," and their children were constrained by West German anti-Nazi legislation to refrain from publicly repeating old Nazi propaganda. There were several waves of such successor organizations, each progressively less authentic. They even appeared in Eastern Europe, following the collapse of the communist empire, consisting mostly of violent, anti-immigrant youth gangs and nativist skinhead bands.

Membership

The focus on fascist movements naturally puts the emphasis on the membership. How large were these movements? What was the typical age of the members? At what point did they join up and why? The largest among them—the PNF and the NSDAP—approached the magical 1-million mark at about the time they came to power or shortly thereafter, but their membership was notoriously volatile, as if motivated only by short-term bouts of passion. After taking power, both tended to absorb other existing groups such as youth organizations or other parties. Fascist movements, like most political parties, were inclined to brag about their membership figures or to keep poor records. Hectic membership growth was also related closely to the postwar demobilization from military service, though fascist movements were in competition with veterans groups and, in Germany and Austria, with large right-wing vigilante organizations (*Einwohnerwehr*). A large influx also occurred after election triumphs and the takeover of power.

At the very beginning, of course, the membership, say of the Italian "fascists of the first hour" or the Nazi "old fighters" was very small. Their first goal, in terms of membership, had to be to acquire a mass following and audience amid the huge followings of their left-wing rivals, given the burgeoning mass politics of most interwar countries.

Social Composition

Although there is little and often unreliable information about the social composition of most interwar European fascist movements, the question of their socioeconomic origins, from their beginnings, took on great significance. Given the rising challenge of labor movements and the prominence of Marxist interpretations in advanced industrial countries throughout much of the 20th century, many analysts equated socialist and communist parties with the rising industrial proletariat while associating fascist movements with the militant defense of the interests of the bourgeois middle class and old aristocratic elites. Fascist antisocialism was thus interpreted as the bourgeois "class struggle *à rebours*" (the class struggle in reverse) against the advance of working-class parties and trade unions.

There was of course much evidence for this antisocialist behavior of fascists, although they themselves resisted such an interpretation and frequently directed their polemics also against capitalism and the "reactionary" social elites. The NSDAP even called itself National Socialist German Workers Party and often pointed out its worker-friendly program. Nearly a third of its rank-and-file membership in 1935, indeed, were blue-collar workers, and one fifth were low-level white-collar employees. To be sure, the German Socialists (*Sozialdemokratische Partei Deutschlands*; SPD) and Communists (*Kommunistische Partei Deutschlands*; KPD) had a far larger working-class component, and the blue-collar share of the total German population at the time was 46.3% (Peter Merkl, 1980a, Tables 3–5, pp. 765–767; Tables 9–11, pp. 774–776). The guiding formula of most contemporary fascist movements was a third-force formula combining strident nationalism with a form of non-Marxist socialism.

But the crux of the matter is that it makes little sense to equate Karl Marx's sweeping concept of the rising proletariat with the fragmented occupational statistics of socialist and Nazi movements. Occupational status does not necessarily imply a class role and ideology. It makes a difference, also, whether a worker is skilled or unskilled, a foreman or just an ordinary worker; whether he or she belongs to a well-established, traditional trade, is a union member or not, and so on. The Nazis and the Italian fascists even had their own, large trade unions and so did various conservative parties, such as the Catholic Center Party in Germany and the Christian trade unions in several other European nations.

The status differences within fascist parties also matter: The blue-collar share among all NSDAP party office holders, for example, was much smaller—only 23% in 1935 (Merkl, 1980a)—than among the rank and file. Neither Hitler nor Mussolini nor comparable European fascist leaders were workers, although at the start of his political career, Mussolini was a revolutionary socialist intellectual but hardly a worker.

Failing to clearly associate Nazis and other fascists, social analysts then tried to concentrate on exactly how the NSDAP differed from the population of Germany in 1935: According to Merkl (1980a), the party had far higher percentages of white-collar employees (20.6%), civil servants (9.6%), teachers (3.4%), professionals (3.2%), business owners (7.5%), and independent artisans (19.5%).

Unable to attribute fascism to bourgeois and upper-class interests alone, the analysts tried to label the NSDAP a "lower-middle-class" or "petty bourgeois" movement. But although, the German white-collar occupations, for example, had grown prodigiously since 1900, the concept of a lower middle class united by a common awareness of its interests is quite vague. Does the professional military, the civil service, or the large farm population belong to it or not? It was never a part of any overarching class struggle, which would have assigned it a role in the dynamic of the larger society.

Change in the Movements

In the dozen or so years since its beginnings, furthermore, the NSDAP changed radically, and the disproportionate numbers of demobilized, militant veterans and university students of the early (Bavarian) phase were replaced by white-collar and blue-collar elements in the party of the 1930s. There were similar groups outside Bavaria before 1925, such as the Freikorps (Free Corps) and the *Deutschvölkische Schutz- und Trutzbund*, but there is a lack of sufficient data for comparison. By the early 1930s, farmers, blue- and white-collar employees, and victims of foreclosure, bankruptcy, and mass unemployment swelled the ranks.

Once in power and able to dispense favors, like jobs to the unemployed, the NSDAP grew prodigiously. Particularly civil servants (earlier barred from joining) and professional and business people in small towns and rural areas joined up so that, by 1937, its membership stood at about 2.5 million. But what had been dangerous rebels before the Beer Hall Putsch of 1923, when their demonstration was shot at by the police, had become a rather tame collection of conformists. Only the leadership, not the movement as such, had taken over the state, which continued to be run by its specialized services. The Italian PNF experienced similar rapid growth after coming into power with the bold March on Rome in 1922 and accumulated well over a million members by 1929, when even the Vatican was ready to conclude a concordat with Mussolini that became the 1929 Lateran Treaty.

Statistical gaps and differences in national population statistics make it harder to compare the social composition of the 20-some other fascist movements of interwar Europe. Many of the smaller ones kept no records, but many had their own historical interpretations of themselves, usually outside the Marxist scheme and not derived from the judgment of their antagonists. As in many revolutionary situations, university students often played a big role among the various fascist movements, whether they were the prewar French *camelots du roi* with their integral nationalism and anti-Semitism, the anti-Russian Finnish Academic Karelian Society (AKS), or the Moldavian student disciples of the anti-Semite Alexander Cuza from whom the Romanian Iron Guard of Corneliu Codreanu eventually sprang with a nativist and frenetically religious fascism found otherwise only among the Belgian Rexists and the Slovakian Hlinka fascists.

Marginal Fascist Movements

Were there other fascist movements, perhaps in India, Argentina, or the Middle East, far from rising Marxist labor parties? For many Balkan or East European successor states, left in the wake of falling empires where power gravitated toward old imperial elites, a Marxist interpretation seems beside the point. Fighting off ethnic secession movements with socialist overtones was not the same as facing the challenge of a rising labor movement. Even the post-Tsarist Whites in the Russian civil war and post-Hapsburg Hungary fit this category: The old aristocratic, landowning elites supported the growth of ex-military and fascist movements against the urban Bolsheviks of St. Petersburg and Budapest and ethnic-nationalist rebellions. The Bolsheviks won in the end in Russia, though hardly because of the support of an industrial working class. The violence of the first war and upheavals involving "dictatorships of workers and soldiers" from Russia and Finland to Budapest, Vienna, Munich, and Berlin made for civil war–like confrontations and attitudes: Revolution and violence create militants.

But the farther west and north one goes, if not always toward greater industrial development, the greater is the role of the Marxist image of capital versus workers and, therefore, the significance of the social composition of self-styled fascist movements. In the newborn Austrian republic, the *Heimwehr* veterans were perhaps more reliable supporters of the aristocrats and landowners of the old order than the German-supported Austrian NSDAP of 1934 and 1938. Its fight against the socialists and organized workers was hardly a defense of capital that could be exhaustively described by the Marxist formula. In Italy, the separate developments in the north—which was closer to the model—and the Nationalists of the south who were eventually co-opted by Mussolini also did not quite add up to a victory of the bourgeoisie in a class struggle *au rebours*.

The role of the Catholic Church in Southeastern and Southern European fascism was nowhere stronger than in the complex fronts of the Spanish Civil War. Spain had not been part of World War I, but the external intervention of General Franco's colonial army in defense of a mix of well-established interests (monarchists, landowners, capital, and the Church) was decisive in beating down the republican coalition, not to mention secessionist challenges. The real fascists, the Falange/JONS, were castrated, forced into a merger with monarchists, and sidelined as a kind of popular ideological auxiliary by the real winner, the military dictator Franco. The resulting Falange/Movimiento became another toothless fascist façade not unlike the Nazi storm troopers (*Sturm Abteilung*, or SA) after 1934.

A Generational Explosion?

These interwar movements mostly started in a grandiose manner and, then, even though they reached amazing numbers, began to decline in militant spirit. Could their age composition explain their counterrevolutionary power? Their earliest years, whether right after the Great War or at the onset of the Great Depression, were often marked by the surprising youth of most of the militants. Fascist leaders, though not particularly young themselves, tended to take advantage of the enthusiasm and energy of the young with slogans like "Move over, you old ones!" or by celebrating youth in their songs—*Giovinezza*, *Giovinezza*. They often pointed to the generational abyss between their own parties and their more established and gerontocratic rivals (except the equally young Communists).

The long period of peaceful prewar development, in particular, and the war itself often led to a kind of sclerosis within progressive movements and then an explosion of generational conflict. Young and old clashed over their goals, which helped produce the fatal wartime splits among socialists or between them and the communists. In some battlegrounds, such as Berlin between 1928 and 1933, the two ferocious young armies—the young Nazi storm troopers and the Communist Red Front Fighters—were mirror images of each other, both from proletarian backgrounds but one nationalistic and the other socialist. Right at the end of the war, Mussolini's *squadristi* boasted as their antecedents the equally young *arditi* shock troops of the Italian army, while the brown-shirted *Ordner* (bouncers) of the NSDAP rallies claimed the youthful *Sturmbatallione* (elite battalions) of the German army.

The actual ages at which would-be fascists were recruited may have been even lower than that of the soldiers: Comparative statistics are limited, but they suggest that nearly half of the Bavarian Nazis of 1923, 37% of the NSDAP of 1930, and one fourth of the Austrian NSDAP of 1934 (the year of the Dollfuss Putsch) were under 24 years of age. One third of Danish, Norwegian, and British fascists of the 1930s were under 26 (Merkl, 1980a, Tables 6–7, pp. 768–770). One empirical study, based on the Abel Collection of autobiographies of early Nazis, emphasizes the large number of German youth groups of the 1920s, many of a political sort from Left to Right, that the young would-be Nazis had briefly joined before homing in on the storm troopers or the party itself. The youngest to join militant Nazi organizations were barely teenagers. Like the neo-Nazi skinheads of 60-some years later, many already had records of criminal violence and confrontations with authority. The very young were also the most violent and the most likely to participate soon in marching in demonstrations and proselytizing for the movement. It is generally accepted that what attracted these youngsters most was largely the sheer impetus of an extremist populist movement and the constant action and comradeship and to a lesser extent the ideological content of what they were fighting, marching, and proselytizing for, including anti-Semitism. There was quite a difference between them and the relatively less violent but also more racist or anti-Semitic adult party members who never joined the storm troopers.

Investigations of contemporary neo-Nazi and skinhead violence thus reveal major differences between them and their interwar predecessors. First of all, there appears to be a broad range of very young perpetrators of antiforeigner violence, including soccer rowdies, among whom real political affiliation with neo-Nazi groups is rare though many wear pins and insignia that suggest otherwise. They are part of today's very violent European youth scene, especially those 15 to 20 years old, and resemble youth gangs and their activities all over the world far more than their counterparts who took part in political violence of the interwar period (see, e.g., Merkl, 1995; Willems, 1993). Older skinheads tend to be more political and often anti-Semitic and affiliated with neo-Nazi parties that vainly struggle to attract the young hooligans. There is little in common with the quasi-military, disciplined organizations of the fascist past that aimed at major political goals such as overturning the peace settlements of World War I.

Peter H. Merkl
University of California, Santa Barbara
Santa Barbara, California, United States

See also Authoritarian Regimes; Fascism; Party Identification; Party Organization; Totalitarian Regimes; Totalitarianism

Further Readings

Bjørgo, T. (1995). *Terror from the extreme Right*. London: Frank Cass.

De Felice, R. (1966). *Mussolini il fascista: la conquista del potere* [Mussolini the fascist: The conquest of the power]. Milan, Italy: Einaudi.

Larsen, S. U., Hagtvet, B., & Myklebust, J. P. (Eds.). (1980). *Who were the fascists: Social roots of European fascism*. Bergen, Norway: Universitetsforlaget.

Merkl, P. H. (1980a). Comparing fascist movements. In S. U. Larsen, B. Hagtvet, & J. P. Myklebust (Eds.), *Who were the fascists: Social roots of European fascism* (pp. 752–783). Bergen, Norway: Universitetsforlaget.

Merkl, P. H. (1980b). *The making of a storm-trooper*. Princeton, NJ: Princeton University Press.

Merkl, P. H. (1995). Are the old Nazis coming back? In P. H. Merkl (Ed.), *The federal republic at 45: Union without unity* (pp. 427–484). New York: New York University Press.

Merkl, P. H., & Weinberg, L. (Eds.). (2003). *Right-wing extremism in the 21st century*. London: Frank Cass.

Payne, S. G. (1976). Fascism in Western Europe. In W. Laqueur (Ed.), *Fascism, a reader's guide: Analysis, interpretations, bibliography* (pp. 295–311). Berkeley: University of California Press.

Ramet, S. (Ed.). (1999). *The radical Right in Central and Eastern Europe since 1989*. University Park: Penn State University Press.

Willems, H. (1993). *Fremdenfeindliche Gewalt: Einstellungen, Täter, Konflikteskalation* [Xenophobic force: Attitudes, culprit, conflict escalation]. Opladen, Germany: Leske & Budrich.

FEDERALISM

Federalism is often referred to in current writings on politics as an old political idea whose time has finally come in the 21st century. In this respect, a distinguished scholar of this subject, Ronald L. Watts (2008), has noted "there are at present some 25 countries encompassing over 40% of the world's population, and each exhibiting the fundamental characteristics of a functioning federation" (p. 1). And he points out that "a distinctive feature about this current popularity of federalism in the world is that the application of the federal idea has taken a great variety of forms" (p. 7).

The initial framing of "federalism" in its modern manifestation, both as an idea and as an institutional form, is generally attributed to two prominent American Constitutional Fathers, Alexander Hamilton and James Madison, the principal authors of *The Federalist*, in October 1787. In promoting this concept, they sought to defend a new hybrid regime form that they and other framers of the American Constitution had proposed at the founding Philadelphia Convention on September 17, 1787, for later ratification. It was intended to strike a balance between the decentralized and centralized forms of confederal and unitary government. The former had been applied to the much more decentralized political union that had operated from 1781 to 1787 in the 13 newly independent colonies of British North America under the Articles of Confederation. In this earlier short-lived and highly unstable political arrangement, the original constituent units were essentially in control of the newly formed and fragile central government. A similar pattern of union had operated in the German, Dutch, and Swiss Confederations of the 16th and 17th centuries. Unitary government, on the other hand, was the pattern of rule that operated in most other states in which political decision-making power and authority were vested in one unit or center. The American Constitutional Fathers favored federalism as an intermediate form between centralized and decentralized government because they believed that by dividing sovereignty in this balanced way, they could prevent the concentration of authority and promote liberty and political pluralism in a new union. They considered federalism to be a vertical extension of the concept of separation of powers between the executive, legislature, and judiciary, which they had adapted from the noted 18th-century French political philosopher, the Baron de Montesquieu.

Defining Principles and Characteristics

The major and defining principle of this new regime form, according to the British constitutional thinker Kenneth C. Wheare in his pioneering study *Federal Government* (1964), consisted of a division of power and authority between two different political decision-making units—a central one exercising jurisdiction over the entire nation-state and a constituent or regional one having responsibility for a part of that nation-state. These two governmental units would be recognized as having equal status in law, so that neither one could abolish or subordinate the other. Other characteristics of federalism, such as the capacity to exercise authority and decision-making power directly on their respective populations and the right of each unit to legislate or regulate policy within its assigned sphere of authority, flowed directly from this principle. The same was true of the various institutional elements or manifestations of this basic idea. These included the institution of two complementary legislative bodies—one to represent the center on the basis of population and the other to represent the constituent units in terms of territoriality—and the establishment of courts as judicial arbiters in jurisdictional disputes between the two levels of government. Both levels would also be expected to participate in a formal and fundamental way in any future constitutional amendments.

Pre-1945 Theoretical Contributions

This general concept of federalism gradually gained prominence and approval by some leading 19th- and early-20th-century political thinkers, including luminaries like Alexis de Tocqueville, Albert V. Dicey, James Bryce, Pierre-Joseph Proudhon, and Harold Laski (Sobei Mogi, 1931). Like the authors of *The Federalist*, they viewed federalism in both normative and descriptive terms largely as an institutional device designed to divide sovereignty and prevent the concentration of authority and power in a single decision-making center, promote political pluralism, and maximize liberty. It was entrenched as a regime form in the 19th and early 20th centuries in several countries besides the United States, including Switzerland, Canada, and Australia. It was also embraced by the new postwar regimes in Germany and Austria and, at least in a formal sense, by a number of newly independent Third World countries, such as Brazil, Argentina, Venezuela, and Mexico in Latin America; South Africa and Nigeria in Africa; and India and Malaysia in Asia. Still, its adoption, both intellectually and institutionally, was limited and slow. Moreover, it faced widespread opposition globally from critics for what was perceived to be its predisposition to cumbersome and inefficient governmental decision-making processes, pervasive internal conflict and division, and regime instability and breakdown.

Post–World War II Theoretical Trends

After World War II, the core ideas and writings on federalism began to diverge in political science and to coalesce around two distinct traditions, which are generally designated as Anglo-American and continental European. The former (the Anglo-American tradition) encouraged the formulation of new analytical approaches to federalism, which paralleled conceptual and paradigmatic changes in the discipline as a whole. The proponents of these new approaches to federalism, like the advocates of the reform-oriented behavioral movement in political science during this period, generally embraced positivist, empirically oriented, systematic, and individualistic intellectual perspectives in the study of politics. They also sought to minimize and downplay the significance of values in their analyses.

The Postwar Anglo-American Tradition

An early and seminal contribution to this postwar tradition was made by Wheare (1964), who defined federalism both constitutionally and in governmental practice as that system of government in which authority is divided between national and regional governments, so that each, within a sphere, is coordinate (i.e., co-equal) and independent. For Wheare, a principal focus of analysis for identifying the degree of federalism that existed in the constitutions and governmental practices of nation-states was the division of legislative powers between its levels of government. Wheare applied this analytical device to a comparison of the institutions and working practices of four "mature federations" of that time: Australia, Canada, Switzerland, and the United States. His approach was rapidly and widely accepted as an appropriate template for identifying and analyzing federalism.

But as occurred in post–World War II political science in general, criticism by federalism specialists began to be directed at what was seen as the excessive formal-legalism of Wheare's approach. Their critique encouraged other Anglo-American scholars to redefine federalism in different terms and contribute to its further theoretical development. They included William S. Livingston, who viewed federalism from a sociological standpoint as "a device by which the federal qualities of a society are articulated and protected" (Livingston, 1952/1971, p. 22); William Riker (1964), who defined federalism in political and rational choice terms as a "bargain"; Carl Friedrich (1968), who envisaged federalism as an ongoing "process"; and Daniel Elazar (1962), who first analyzed federalism as a cooperative "partnership" between the two major levels of government. Later, in conjunction with Vincent Ostrom (1970), he redefined the concept of partnership as a "covenant" or legal and moral contract between these governmental partners (Elazar, 1987).

There were also some strong criticisms directed at these new postwar approaches to federalism by William Riker and S. Rufus Davis. Riker questioned whether federalism as a regime form even existed and, if it did, whether it mattered at all (Riker, 1969). Davis attacked the post–World War II predilection for redefining and applying different analytical constructs of federalism as an artificial doctoring of what he considered to be an ambiguous and value-laden concept consisting of many different, noncomparable concrete forms in different societal settings (Davis, 1978). After this rather fertile period in the theoretical evolution of federalism from the late 1940s to the 1970s, there was a dearth of innovative and influential Anglo-American contributions to federal theory.

The Postwar Continental European Tradition

Early Mentors

The proponents of the postwar continental European tradition of federalism generally looked to different mentors or guides for their intellectual inspiration (Michael Burgess, 2000). For many of them, the originators and intellectual founders of modern federalism were not the principal authors of *The Federalist*. Rather, they considered themselves to be the theoretical disciples of Johannes Althusius, a German, Protestant (Reformed) local councilor from the city of Endic, who remained a relatively obscure social and political thinker during his lifetime and for the ensuing two and a half centuries. Even though he lived and wrote almost 200 years prior to the drafting of the American constitution, Althusius's ideas are often regarded today as a bridge between premodern and postmodern Western political thought. More precisely, they are viewed as a precursor of the pluralist critique of statist and sovereigntist ideas that have dominated Western thinking since the establishment of the Westphalian state system in the mid-17th century. He viewed federalism in broad terms as a manifestation of plural governance, popular consent, and a normative commitment to social solidarity. His concept of federalism was multilevel and societal rather than two-tiered and narrowly political. It was derived from his understanding of politics as a "consociation" or social network of generically equivalent associations linking family and kinship groups to social and economic guilds or colleges, cities, rural districts and provinces, and ultimately a universal commonwealth (Thomas Hueglin, 1999, pp. 114–116).

Another early inspiration for this postwar continental tradition of federalism can be found in *Du principe federatif* (*The Federal Principle*) published in 1863, one of the last works of the influential 19th-century anarcho-socialist French philosopher Pierre-Joseph Proudhon. Like Althusius, he advanced a notion of federalism embodying an organic view of state and society, in which emphasis is placed on the whole person in social interaction with others, rather than on the more socially isolated individual. He likewise incorporated into his federal idea the medieval concepts of corporatism (i.e., representation of functional associations in public policy making) and subsidiarity (i.e., investing decision-making authority in a governing unit closest and most attuned to the needs of the mass of the population, wherever feasible). Also, he espoused a normative and secular rather than religious conception of federalism.

Proudhon's federalist thought also resembled that of Althusius in that he viewed federalism as a series of agreements struck by the associational representatives of socially engaged individuals. These associations were likewise multitiered, linking the smallest and most primary group, the family, to the largest and most remote of secondary associations, the state. But unlike in Althusius's vision, his corporatism embodied the free association of economic producers and workers interacting in what he considered to be the socialist ideal of a "mutualist" society. In this type of socioeconomic order, goods are exchanged between individuals on the basis of a "just value" defined in terms of labor input.

Both of these earlier philosophers had a significant impact on post–World War II continental European intellectuals and political activists who advocated the formation of some type of loose supranational federal structure for Europe. This was also the case for some leading social Catholic thinkers of late-19th- and early-20th-century Europe, who derived many of their ideas from the papal encyclicals *Rerum Novarum* (1891) and *Quadragesimo Anno* (1931). While generally omitting any direct reference to the concept of federalism itself, they helped popularize some of the central concepts previously identified with the early antecedents of continental European federal thought. These included an organic view of state and society based on corporatism and subsidiarity, an emphasis on the socially interactive person rather than the atomized individual, and the fostering of a normative view of politics based on ethical-religious principles.

Postwar Manifestations

There have been two distinct post–World War II manifestations of the continental European federal tradition. The first was that of the integral federalists, a mixture of institutionally unattached intellectuals and political activists, who advocated a so-called maximalist revolutionary concept of federalism for postwar Europe. The second is a more mainstream and traditional intellectual group made up primarily of university-affiliated social and political scientists, who promote a more flexible, inclusive, and looser form of federalism for the postwar European Union (EU), which is widely referred to as multilevel governance (MLG).

Integral federalism was initially established in France in the 1930s by a group of continental European thinkers who called for a pan-European federalism based on Proudhonian principles, including antinationalism, decentralization, and mutualism, designed to combat the growing conflict between France and Germany at that time. However, it was forced to dissolve in 1939 in the face of the looming global conflict. Its original members regrouped and formed a more broadly based intellectual and political group in 1944 that included some prominent social Catholic thinkers who were in the vanguard of the movement for a pan-EU in the post–World War II period. Their ideas included reorganization of European society from the ground up according to organic and corporatist principles, transnational integration of different branches of the economy as a precondition for the creation of supranational economic structures, and the establishment of new political structures that involved regional decentralization, devolution of power to functional economic agencies, and the creation of a constituent assembly to fashion a new pan-European legal-political order along federalist lines. Although the integral federalists attempted to give these ideas more concrete meaning by applying them to the particular conditions of postwar Europe, most advocates of a supranational European community rejected their views as too radical and utopian and as lacking in an accompanying strategy for their practical realization. As a result, the direct influence of the integral federalists on the evolution of the European Community began to wane by the late 1940s. Since then, the handful of remaining members have focused more on political education than on practical politics (Lutz Roemheld, 1990).

Those who are committed to applying the concept of MLG to the political analysis of the EU today constitute the most recent and probably the most influential global proponents of new federal theoretical ideas. They comprise, primarily, university-based social and political analysts. They first began to promote the use of the concept of MLG rather than federalism to describe the EU in response to what they considered to be the unique character of the new supranational governance form that had emerged after the passage of the Single European Act in 1985 and the negotiation and implementation of the Maastricht Treaty in the early 1990s. Their writings on MLG have been numerous and wide-ranging, and their definitions of the concept, although manifesting some common or shared characteristics, are highly divergent. They have sought to describe and analyze the new multilayered European political entity consisting of overlapping jurisdictions by first broadening the idea of federalism to encompass more than two levels of government and more than two autonomous policy-making structures; they now include both regional/supranational and local government tiers. Their concepts have other common strands—most notably, accepting or even promoting the participation of nonstate actors, such as nongovernmental organizations (NGOs), corporations, and unions, in intergovernmental decision making; observing and applauding the proliferation of overlapping decision-making networks engaged in these functions; pointing to the change in the style of state decision making from that of command and control to steering, coordination, and networking; and exposing new and difficult challenges facing MLG structures, such as how to assign governing responsibility and how to foster democratic accountability (Ian Bache & Matthew Flinders, 2004).

The concept of MLG has since been extended by its proponents to other areas of systematic political analysis besides the decision-making structures and functions of the EU, among them regionalism and urban governance. It has also gained followers in other continents and geographic areas, including North America, South America, Asia, and Africa, and it is increasingly applied as a tool of political analysis, largely due to its flexibility and adaptability to the existing but rapidly changing global economic and political conditions.

However, its adherents have also been subjected to a broad range of criticisms, including that of being too descriptive and insufficiently predictive; exaggerating the importance of subnational actors in decision making; neglecting the implementation and outcome stage of policy making; overemphasizing what they describe as the postconstitutional and extraconstitutional nature of the patterns of governance that they try to encompass; giving priority to accommodation, consensus, and efficiency in governance over the core values of democratic government; and above all, acceding too readily to definitional imprecision and conceptual stretching by applying MLG indiscriminately "to any complex and multifaceted political process" (Bache, 1998, pp. 153–154).

Since the mid-1990s, there has been an increasing blurring of the distinction between the concepts of federalism and MLG, again primarily as a result of the growing complexity in patterns of intergovernmental relations in federal systems emanating from the forces of globalization. The historical and analytical relationship between these two concepts tends to operate in a multidirectional, interactive, and interdependent manner, so that their respective theoretical developments are mutually intertwined and overlapping in both content and direction. Thus, MLG may be regarded as a useful addition to the conceptual toolbox of scholars of federalism.

Future Theoretical Directions and Applications

If the concept of federalism continues to evolve in this flexible manner in the face of intensifying globalization, it is likely that it will be increasingly embraced in positive terms by both political analysts and practitioners across the globe. This is particularly the case if the two post–World War II theoretical traditions of federalism—the Anglo-American and Continental European—persist in their current patterns of mutual interaction, influence, and convergence. And it is also likely to manifest itself in future political studies that draw on either or both of the prevailing definitions of federalism and MLG.

Whether the same pattern of convergence will develop in future institutional applications of these two concepts is more difficult to predict. A few developed and developing countries, such as Spain, South Africa, Russia, Ethiopia, and Nigeria, are currently experimenting with or are still considering using some type of federal institutional form for their regimes. Others, such as Yugoslavia, Czechoslovakia, the Sudan, Iraq, and Sri Lanka, have previously embraced or have seriously contemplated adopting federal structures and found them to be inapplicable to conditions in their countries. Therefore, it is uncertain whether the recent global trend toward expanding federalism as an institutional device in the 21st century will persist. But there is little doubt that this concept, in one form or another, will continue to influence in significant ways the political thought and analysis of future generations of political scientists.

Michael B. Stein
University of Toronto
Toronto, Ontario, Canada

See also Pluralism

Further Readings

Bache, I. (1998). *The politics of European Union regional policy: Multi-level governance or flexible gatekeeping?* Sheffield, UK: Sheffield Academic Press.

Bache, I., & Flinders, M. (Eds.). (2004). *Multi-level governance.* Oxford, UK: Oxford University Press.

Burgess, M. (2000). *Federalism and European Union: The building of Europe, 1950–2000*. London: Routledge.

Davis, S. R. (1978). *The federal principle.* Berkeley: University of California Press.

Elazar, D. J. (1962). *The American partnership: Intergovernmental cooperation in the nineteenth century United States*. Chicago: University of Chicago Press.

Elazar, D. J. (1987). *Exploring federalism.* Tuscaloosa: University of Alabama Press.

Friedrich, C. J. (1968). *Trends of federalism in theory and practice.* New York: Praeger.

Hueglin, T. O. (1999). *Early modern concepts for a late modern world: Althusius on community and federalism*. Waterloo, ON, Canada: Wilfrid Laurier Press.

Livingston, W. S. (1971). A note on the nature of federalism. In J. P. Meekison (Ed.), *Canadian federalism: Myth or reality* (2nd ed.). Toronto, ON, Canada: Methuen. (Reprinted from *Political Science Quarterly*, *67*(1), 81–95, by W. S. Livingston, 1952)

Mogi, S. (1931). *The problem of federalism: A study in the history of political theory* (Vols. 1–2). London: Allen & Unwin.

Ostrom, V. (1970). *The political theory of a compound republic: Designing the American experiment*. Lincoln: University of Nebraska Press.

Riker, W. H. (1964). *Federalism: Origin, operation, significance.* Boston: Little, Brown.

Riker, W. H. (1969). Six books in search of a subject or does federalism exist and does it matter? *Comparative Politics, 2*(1), 135–147.

Roemheld, L. (1990). *Integral federalism.* New York: Peter Lang.

Stein, M. B. (1968). Federal political systems and federal societies. *World Politics, 20*(4), 721–747.

Watts, R. L. (2008). *Comparing federal systems* (3rd ed.). Montreal, QC, Canada: McGill-Queen's University Press.

Wheare, K. C. (1964). *Federal government* (4th ed.). New York: Oxford University Press.

FEMINISM

The 20th century was a momentous one for feminism and politics. Beginning with the spread of women's voting rights, the political participation and engagement of women expanded during the decades that followed. Links were forged between women's workforce participation and their active involvement in public affairs. Women's civic associational activity, often voluntary, brought the public and private worlds of citizenship together. Women's liberation movements reached across age, class, and racial barriers to forge new policy demands. Gender inequalities in representation, too, mobilized coalitions of women determined to open up formal politics. Women played active roles in the democracy-seeking political movements of the latter part of the century. Although the adoption of feminist insights for making sense of the world of politics was relatively slow in the field of political studies, gender and politics scholarship is now a coherent subfield within the wider discipline. Over the past few years, this body of scholarly research has significantly evolved. New research questions, concepts, theories, and methods have emerged in response to new developments in the feminist struggle for women's empowerment, together with important developments in the social sciences and in the wider context of world politics. Against this background, this entry dwells on the legacy of the 20th century for women's advancement in politics, maps out the main methodological approaches used by feminist political scientists, and outlines the main thematic areas in the study of politics and gender.

Women and Politics: The Legacy of the 20th Century

In 1906, Finland became the first country in the world to give women the right to vote and to stand for election. One year later, 19 women were sent to the Finnish parliament. This is a momentous landmark in the story of women's participation in public affairs since it was the first time in history when involvement in democratic politics became a right for all women and not the privilege of some. In addition, this also represents a first milestone in a century of unprecedented activism on gender equality.

The right to vote and be elected to parliament was an agenda around which the first great wave of women's mobilization rallied—a movement that soon became a global one as it spread rapidly, with mass demonstrations, rallies, and marches. Winning the vote created a demand for women to become political party activists, speakers at public meetings, and candidates, though their latter role was largely tokenist at the time. It realized, in some measure, the dreams of 18th- and 19th-century feminist political philosophers and women's rights advocates such as Olympe de Gouge, Mary Woolstonecraft, and Harriet Taylor Mill. However, not all countries accorded women the right to vote at the turn of the 20th century as in many of them, women did not have access to this political right until well into the middle of the century. Thus, Spain accorded women voting rights in 1931, France in 1944, and Italy in 1946. Indonesian women won the vote in 1945, while Peruvian women followed a decade later in 1955. Switzerland finally extended the franchise to women in 1971, while the Kuwaiti parliament passed a law enfranchising women in 2005.

As the century moved on, women's sense of civic identity became linked with their ambitions for personal fulfillment. The years of World War II were defining in this regard as the armaments industry, and the various services supporting the war effort drew increasing numbers of women into the workplace. War-induced employment, presented as a patriotic act, provided low-paid women already in work with better jobs and conditions and brought married women into the workforce (although most women of working age remained in the home). After the war, working women either returned to the home or were squeezed back into lower paid positions. Women's status declined, the gender wage gap increased, and as Wilma Rule noted (1974, p. 1715) "many left [employment] voluntarily in the ideological milieu of a new familialism buttressed by a popularized Freudianism which portrayed the non-homebound woman as psychologically afflicted." The fact that women in the 1950s were expected by society to marry and be full-time homemakers means that they lost out on the orientation to civic and public affairs that the world of employment tends to offer, although it did not mean that they were not civically engaged. On the contrary, women's associational life during the 1950s was, by all accounts, active and vibrant, as hundreds of thousands of home-based women joined locally based voluntary women's organizations in Britain and the United States. Yet because these organizations focused on church and charity-related volunteering, improving homemaking skills, civic engagement, and education, associational activity was engaged in for its own sake, not as a political enterprise, although some scholars have suggested that by encouraging the citizen housewife, these associational groups contributed to the blurring of the distinction between the public and the private worlds of women.

By the 1960s, women's membership in these associations had suffered a severe decline. The decade of housewives was followed by the decade of women's liberationists. Many of these women's associations became redefined as locally based civic bodies with national and international advocacy roles coordinated by a professionalized leadership. This is a decade that ushered in a new phase of economic and social development in industrialized societies, which in turn had an impact on women's social roles. Women's educational attainment, an important contributor to social and civic empowerment, had risen and, in the United States and Europe, began to outstrip that of men. Employment opportunities, a technological revolution that drove a new phase in consumerism and communications, and the expansion of middle-class aspirations began to chip away at traditional gender role contracts in both public and private spheres. The antiwar, civil rights, and student protest movements of the 1960s in advanced democracies provided the impulse for the second wave of feminism. The popularizing of authors such as Betty Friedan, Simone de Beauvoir, and Germaine Greer through the new media of television brought the issue of women's oppression into homes across industrialized postwar democracies. At around the same time, legislation in the United Kingdom and the United States had begun to tackle pay and other discriminations against women in the workplace. Government reports in many countries documented women's disadvantaged legal and social status and made recommendations for reform. Women's liberation movements in Britain, the United States, and other countries, engaged in a period of intense activity. Regional and national meetings that attracted hundreds of women, local consciousness-raising groups, marches on parliament, and other high-visibility protests characterized the movement. Connections were made across class and race lines, and for a brief time, all women were united in their commitment to end discrimination. The sisterly solidarity fragmented shortly afterward, with class, race, and other deep-seated social and political divisions reasserting themselves. The lasting legacy of the women's movement, though, was to change social attitudes toward women's role and status. It had a long-term effect on public policy, on sociocultural norms, and on women's aspirations. What was just as important was that it mobilized and engaged a generation of women and affected the attitudes of a generation of men.

Following the women's liberation movement of the 1960s and 1970s, the last part of the century saw the global diffusion of gender equality norms. This was supported by the United Nations (UN) Commission on the Status of Women, which in turn is dedicated to gender equality and the advancement of women. Responding to the momentum generated by the women's movement in the 1970s, the General Assembly declared 1975 as International Women's Year and organized the first World Conference on Women in Mexico City, followed by others in Copenhagen (1980), Nairobi (1985), and Beijing (1995). In 1979, the UN adopted the Convention on the Elimination of Discrimination Against Women, seen as the international bill of women's human rights. In 1995, it adopted the Beijing commitments and Platform for Action to secure gender equality and peace, and in 2000, it adopted Resolution 1325 on women, peace, and security, calling for women's inclusion in peace-building processes. In 2009, it added to this resolution by explicitly mandating peacekeeping forces to protect women and children from sexual violence during armed conflict.

None of these policies and events would have occurred in the absence of a concerted, focused, transnational impetus from women. Transnational activism was the chief characteristic of women's movements of the end of the century. It linked local issues with global concerns, built connections across borders, and created dialogue between feminists in the Global North and the Global South. It began in the mid-1980s, as preparations for the Nairobi World Conference on Women got under way, although it took some time before women from the Global North and Global South could agree on a common agenda. Two paradigm shifts in the economic and cultural world order prompted consensus building among First World and Third World women: the assertion of transnational neo-liberal markets dependent on cheap labor (often provided by women) and the rise of Islamic fundamentalism. Both were seen as global threats to women's equality and empowerment. A more united transnational feminist activism quickly proliferated, particularly after the Beijing conference. In parallel with this transnational civic and political engagement of women, awareness of the absence of women from elite decision making grew. The importance of their presence in this arena for bringing about gender equality and for representing women's concerns resulted in the widespread adoption of gender quotas in political and public life. By 2000, many European and Latin American countries had adopted some form of gender quota, along with a range of countries in Asia, largely due to women's engagement with party and political systems. However, even currently, women's equality with men in political decision making has some considerable way to go. Their representation stood at 18% worldwide at the end of the century.

Feminism in Political Science: Perspectives and Methods

Despite the important achievements in women's political empowerment that took place over the course of the 20th century, the incorporation of feminist perspectives into the discipline of political science was a much slower process. Thus, it was not until the 1980s that research on gender and politics took off and not until the 1990s that it consolidated itself as a subfield within the wider discipline of politics—with dedicated journals, research networks, projects and events, specialized research centers and institutions, postgraduate programs, and so on. However, to claim that gender and politics has become a coherent subfield does not mean that it is a uniform subfield—quite the opposite, in fact, as we can find a diversity of research agendas that are informed by manifold theoretical perspectives and methods and that are as diverse as the feminisms on which they draw. This section provides a summary overview of the variety of feminist perspectives underpinning most of the gender and politics research that is being conducted today.

Studies on gender and politics are framed within one or more of the main feminist perspectives that have been developed up to the present. These are liberal, radical, poststructuralist, and postcolonial feminisms. Categorizing studies into one or another perspective is not always straightforward since these perspectives are used to address a variety of questions specifically concerned with gender and politics; many opt to develop a modified version of a particular one while incorporating valuable insights from the others.

Liberal Feminism

The first perspective is liberal feminism. Research on gender and politics informed by this perspective is characterized by a number of features. The first one is that the structures of the liberal state are assumed rather than problematized. In other words, the problem is not the structures of the liberal state but the lack of women and women's representatives within its structures. Therefore, for gender and politics scholars influenced by liberal feminism, the task of feminizing politics consists in bringing more women into the political realm. It is a reformist, rather than a transformative, agenda. The second feature of gender and politics research influence by this perspective is its acceptance of mainstream understandings of the political as the world of government and public affairs. Consequently, their subject matter is constructed on a strict distinction between the public and the private realms, where the latter is considered to fall outside the scope of research and, therefore, to be irrelevant to the central research questions. The third feature of this research relates to the methodologies adopted, as it operates with a sharp distinction between fact and value statements, a strong reliance on positivist epistemologies, and a favoring of quantitative methods. Finally, the fourth feature of this research is that it is, to a large extent, focused on individuals and agency for understanding our social reality (e.g., lack of individual women in parliaments) and as a main source of social change (e.g., women political leaders, women's and feminist interest groups).

Radical Feminism

The second perspective is radical feminism. A key feature of gender and politics research influenced by this perspective is its questioning of the public–private split that has traditionally defined the feminist study of political science. The blurring of the boundaries between the public and the private derives from one of the central premises of radical feminism—namely, the idea that women have been traditionally oppressed by a power system of male supremacy (patriarchy) that permeates all social structures and relations. This male hegemony not only applies in public institutions but also in those areas belonging to what has been traditionally understood as the private sphere (i.e., the body, sexuality and reproduction, nurturing and intimacy, the emotions, etc.). Consequently, according to this perspective, there is nothing in the entire social system that is exempted from becoming an object of political struggle. In contrast to research influenced by liberal feminism, gender and politics scholars influenced by radical feminism do not see the project of bringing more women into state structures (legislatures, executives, and judiciaries) as emancipatory unless it is accompanied by other changes that imply a deeper transformation of those existing structures themselves. Another characteristic feature of this research is a shift of attention from the concept of sex to that of gender. Thus, instead of focusing on individual women (or the lack of them) in political institutions, this research is mainly interested in the gendered nature of social structures, including formal political institutions. Finally, in relation to the methodologies adopted by this kind of research, it should be noted that although scholars influenced by the tenets of radical feminism may be suspicious of positivism, no specific feminist methodology has been developed for the study of the political world. Instead, these researchers opt for a diversity of qualitative methods such as discourse analysis, ethnographic analysis, and case studies.

Poststructuralism

The third feminist perspective that has influenced gender and politics scholars is poststructuralism. A distinctive feature of this strand of research is the idea that traditional concepts in this field of research—such as gender, representation, or the state—are discursive constructions rather than objectively or structurally given realities. Using this perspective, the main focus of the gender and politics researcher is not on individual women and their lack of political representation and participation or on social structures and the gender–power relations embedded in formal and informal political institutions but, rather, on the discourses and practices that construct those concepts. A second feature of this research is that it operates with an idea of power understood as a relational, rather than a static, concept. Hence, in contrast to an idea of power as something that individual men or patriarchal structures own and try to retain, from a feminist poststructuralist perspective, power is understood as an unstable and ever-shifting expression of relations in all aspects of every life. A third distinctive feature of gender and politics research influenced by poststructuralism is the abandonment of a normative standpoint from which value judgments are made to appraise or condemn a reality that is taken as objectively given, such as the patriarchal or the women-friendly state. In other words, the aim of research informed by this perspective is not to provide answers but to show how gender power is constructed through relational discursive practices in specific social and political contexts. For example, recent studies on gender and the state argue that the idea that the state is discursively constructed can explain not only why there are patriarchal states as well as women-friendly states but also why it is possible to find significant differences in this regard between different parts of the same state. Poststructural feminist perspectives have also influenced recent work on gender and political representation.

Postcolonial Feminism

The postcolonial feminist perspective (mainly in the field of international relations) challenges the epistemological and ontological bases of liberal and radical feminist research in two important ways. First, it argues that gender oppression is perpetrated by colonialism as well as patriarchy and that the combination of both forms of oppression shape women's roles and their status in non-Western societies. It is critical of the universalist assumptions of women's shared experiences that liberal and radical feminism presume. This leads postcolonial scholars to question the Western discursive construction of Third World women as a singular entity, its inbuilt tendency to homogenize women's experiences, and its unstated valuing of Western women's empowerment. Second, it challenges the inherent racism in the colonial project in historical and contemporary contexts. Postcolonial feminism as a research approach uses the deconstruction of the Euro/Western-centric legacy to recognize the multiple forms of women's activism and empowerment in Third World societies. More recent studies have sought to uncover the shared sites of oppression among women—such as global capitalism—while recognizing the diversity of their traditions, standpoints, and experiences. Postcolonial feminist political research has much in common with postmodern studies: Both seek to uncover the multiple layers of female oppression that are created through discourse and narrative. Both are attentive to the intersections of race, class, gender, and sexuality with national, cultural, political, and other identities.

Established and Emerging Research Agendas in Feminism and Politics

The brief reviews above of women's political empowerment and scholarly approaches to studying gender and politics reveal a close connection between the worlds of activism and research. This section outlines some important thematic areas of research that unite the study of feminism and politics.

The first theme is one that is fundamental to political science—democracy and political representation. Feminist scholars have paid considerable attention to this theme, grounding their work in case and comparative studies of women's political participation and legislative seat holding and electoral and party politics. This research also encompasses extensive studies on gender differences in political attitudes and behavior. An extension of this research examines women's contributions to democracy and democratization processes. The cultural turn in feminist studies is also reflected in this scholarship, with a focus on the role of the media in shaping gender perceptions of politics. The studies in this thematic area are generally conducted within the tradition of liberal feminism and employ positivist and historical methods of scholarship. The concern in this field of research is to identify and explore women's participation in civic and political activity and the sites and issues on which women challenge the state and seek to represent interests and perspectives that may be unacknowledged in the existing political order.

A related thematic research field is that of the relationship between gender politics and the state. This disparate area addresses the many and varied sites of feminist political contestation, including women's movements, nationalism and feminism, policy issues, gender mainstreaming, and the male-gendered nature of political institutions. There are two common characteristics of this research, especially among most recent scholarship. One is a shift of attention from women in politics to gender in politics. This highlights the way in which political institutions are gendered and how they act to construct and reinforce gender–power relations, both in the political world and in the wider society. The second characteristic is the marked influence on gender and politics research of the institutional turn that has taken place in the wider field of political science. Emerging research agendas in this area focus on changes in the modern nation-state with a view to exploring the differential impact of a variety of state architectures (federalism, regionalism, etc.) on gender relations as well as examining how recent trends in globalization, supranationalization, and decentralization call into question established assumptions about gender politics and the state. Again, producing a prolific output of scholarly analysis, this body of work uses mainly qualitative methodologies in both liberal and radical feminist traditions.

Women's advocacy in transnational and global civil society constitutes a more recent thematic area in feminism and politics research. Prompted by observations on the globalizing nature of political life in the 21st century and its effects on women's opportunities for civic and political expression, this scholarship examines the gendered effects of the global political economy and international political institutions. This research strand is firmly grounded in feminist international relations research, where questions of transnational activism and the gendered nature of international political institutions, linkages, and practices are opened to critical review. This thematic field offers the most eclectic use of methodologies in conducting research, and the influence of critical social and political theory on scholarship is apparent.

Yvonne Galligan and Sara Clavero
Queen's University Belfast
Belfast, United Kingdom

See also Democracy, Direct; Democracy, Quality; Democracy, Theories of; Democracy, Types of; Interest Groups; Participation; Representation; Social Movements; State

Further Readings

Basu, A. (Ed.). (2010). *Women's movements in the global era: The power of local feminisms*. Boulder, CO: Westview.

Gelb, J., & Palley, M. L. (Eds.). (2009). *Women and politics around the world: A comparative history and survey*. Santa Barbara, CA: ABC-CLIO.

Lovenduski, J. (Ed.). (2000). *Feminism and politics* (Vols. 1–2). Aldershot, UK: Ashgate.

Rai, S., & Waylen, G. (2008). *Global governance: Feminist perspectives*. Houndmills, UK: Palgrave Macmillan.

Rule, W. (1974). Political implications of gender roles: A review of the literature. *American Political Science Review, 68,* 1706–1723.

FEMINIST MOVEMENTS

Feminist movements have become ubiquitous in the modern world over the last half century, and today, there are few countries in which such organizations do not exist. Feminist movements have mobilized on a vast number of issues ranging from civil and equal rights to violence against women and reproductive rights. They have also organized around issues of war and peace, the environment, racism, and world poverty. Feminist movements were part of the struggles for national liberation in the 20th century and played an important part in recent transitions to democracy. They have always had international connections, but since the United Nations (UN) Women's World Conferences at Nairobi (1986) and Beijing (1995) and the UN World Population Conference at Cairo (1994), there has been an exponential growth of transnational women's groups. In their aims, feminist movements have never been solely concerned with political change; sociocultural change has always been seen as essential to challenging gender hierarchies.

Given the plethora of issues on which women have mobilized, distinguishing feminist movements from women's movements in general has been contentious in the literature. Women's movements mobilize women *as women*, as a constituency, regardless of the issue that they address. The term *feminist movement* is best reserved for those women's movements challenging gender hierarchies that subordinate women to men. This debate has been fueled by historical records showing that both have often intermingled and influenced each other. In recent studies, the definition of feminist and women's movements is now regarded as an empirical question, to be answered by contextual analysis of the discourses and activities of the movements in question. Among scholars, it has become common, given the varieties in feminism, to speak of feminist movements in the plural.

The term *feminism* was first used in 19th-century Europe, but the discourse itself can be traced to the early 15th century. Its principal thesis is that the condition of women is not ordained by nature or supernatural order but can and should be changed. During the era of colonization, Western feminist ideas became familiar in the East and the South, which led to early reforms of women's rights in the first half of the 20th century in a number of countries as diverse as China, Turkey, and much of Latin America as well as in North America and Europe.

Up to the 1980s, women's movements in the South had been critical of the concept of feminism, regarding it as a limited Western approach to women's issues and arguing that social-economic development is of greater importance for improving women's status in the developing world. The critique was shared by many from the former communist world who depicted feminism as a bourgeois ideology and claimed women's equality had been achieved under communism. However, autonomous feminist movements have since then emerged in many countries in the South and East, with their own feminist discourses and agendas of gender justice. The increasing international engagement of feminist movements has led to a consensus in the 2000s on a broad social feminism engaging with global inequalities and social justice as well as including reproductive rights and violence against women.

Social movement scholars have argued that the emergence of social movements is tied to the development of the national state, which created a public sphere for addressing grievances and protest.

Mobilization occurs in periods of crisis, nearly always generating feminist-inspired protest alongside other social movements. In the literature on feminist movements, these periods of heightened activities have been denoted by the metaphor of *waves*. Usually the term *first wave* is reserved for the feminist movements that arose in the second half of the 19th century, the period of industrialization in Europe and Northern America, and the *second wave* for those originating in the major social changes of the 1960s and 1970s. Historians have made a strong case, however, for earlier waves of feminist activity in the period of the English Civil War and the French Enlightenment as well as during the French Revolution, when numerous treatises and pamphlets on the status of women were published. The metaphor is also somewhat misleading as there is ample evidence of sustained contention by women between the waves. Moreover, the timing of protests is not synchronical across the world. Women's suffrage and changes in patriarchal family law were gained in the 1920s and 1930s in Latin America, while in many European countries, such reforms had taken place in the previous decades.

The early social movements of the mid-19th century are the precursors of feminist movements in the Western world. These were the antislavery (abolitionist) and temperance movements in the United States and Britain, and many social reform and religious revival movements in Europe. By the late 19th century, feminist movements had developed, either arguing that women were equal to men, and therefore were entitled to equal rights, or different from men, requiring the recognition of women's unique contribution to society. The first wave has often been reduced to the struggle for suffrage, but it had a much broader agenda of social and legal rights. The campaign for suffrage emerged in the late 19th century; today, the only countries where women cannot vote on a par with men are several Islamic states in the Middle East.

Other issues of the first wave were the reform of family law, access to education, and the right to work. The first wave retained its social reform character, organizing on issues as diverse as public health, social work, temperance, peace, prostitution, and the trafficking of women. Many feminists also campaigned against informal social codes limiting women's activities, such as dress reform, access to public spaces, and bars on female sports. Antiwar activism and campaigns against trafficking led to the first transnational feminist organizations in the 1890s. Often, it is held that the first wave ended in the West after World War I, but in many countries the feminist movement remained active in social reform issues beneficial to women, such as maternal leave. Such demands were able to cross class divides. One of the major contentious issues has always been that of how far feminist movements were speaking on behalf of all women and not just for the middle class and whether working class women stood to gain by suffrage and equal rights.

Demobilization occurred in the 1940s and 1950s, but the revival of feminism owes much to entrenched feminist networks and women's institutions during this period. The second wave emerged in the wake of the early protest movements of the 1960s, such as the civil rights, student, and antiwar movements in the West and the rising tide of national independence and social revolutionary movements in Latin America and the decolonizing world. Often called the new or the autonomous women's movement, its emergence galvanized existing women's organizations across much of the world. Differences in organizational style, action repertoire, and ideology distinguished the second wave from the older branches of feminisms in many countries. The new feminist groups were highly critical of formal politics, organized in loose networks, and used street protest instead of the traditional lobby to reach their goals. They created a new radical feminist ideology that emphasized consciousness raising and developed a far-reaching critique of patriarchy, culture, and men.

Politicizing the body, by setting abortion and reproductive rights, sexuality (including lesbianism), and (sexual) violence on the political agenda, also originated in the new feminist groups. These issues have become the distinctive marks of second-wave feminism, leading to worldwide reforms and to strong opposition from (religious) conservatives in general. But the second wave also set out to achieve time-honored goals, such as equal political representation, access to education and work, and decent living conditions in the developing world. The creation of a vibrant women's culture, with its own media, women's centers, and networks, has also been integral to most feminist movements. In countries faced with authoritarian or military regimes, a constituency of women originated who fought for democratization and rights for women in new constitutions, as in South Africa and many Latin American states. This was much less the case in the former communist states, where feminism became discredited by its connection to the former regimes.

One of the crucial issues the feminist movement has been confronting is the diversity of women: No organization can claim to speak on behalf of all women. Differences of class, ethnicity, sexual orientation, age, and disabilities have challenged unity and the choice of priorities. These debates occurred in different sequences in most countries and in the international arena. Scholars generally agree that the feminist movement is one of the most successful of the 1960s social movements and the one with the greatest longevity. This is testimony to the increasing success of feminist movements in bridging diverse communities and coping with the changing global environment. Scholars disagree about the end of the second wave. Some scholars have contended that the new autonomous movement disappeared by the early 1980s. Others have pointed out that many second-wave groups have turned into professional lobby organizations still targeting government and parliament on women's issues into the 2000s, a phenomenon also well documented at the supranational level.

The study of feminist and women's movements, like social movements in general, has fallen victim to the division of labor between political scientists, historians, and sociologists. Political science, with its strong distinction between the public and the private, has emphasized formal institutions and political behavior. Sociology tackled social movements but has often paid only nominal attention to feminist and women's movements, raising questions about the validity of its theories. Women's studies and women's history, both legacies of second-wave feminism, have produced most of the current knowledge; as of now, these findings have been only incompletely incorporated in the other disciplines, and they do provide the most promising material for further study.

Joyce Outshoorn
Leiden University
Leiden, Netherlands

See also Feminism; Gender; Rights; Social Movements

Further Readings

Akkerman, T., & Stuurman, S. (Eds.). (1998). *Perspectives on feminist political thought in European history: From the Middle Ages to the present*. London: Routledge.

Ferree, M. M., & Mueller, C. M. (2004). Feminism and the women's movement: A global perspective. In D. A. Snow, S. Soule, & H. Kriesi (Eds.), *The Blackwell companion to social movements* (pp. 576–608). Oxford, UK: Blackwell.

Ferree, M. M., & Tripp, A. M. (Eds.). (2006). *Global feminism: Transnational women's activism, organizing and human rights*. New York: New York University Press.

Nelson, B., & Chowdhury, N. (Eds.). (1994). *Women and politics worldwide*. New Haven, CT: Yale University Press.

Waylen, G. (2007). *Engendering transitions: Women's mobilization, institutions and gender outcomes*. Oxford, UK: Oxford University Press.

Feminist Theory in International Relations

Feminist theory has developed in the context of the feminist movement and seeks to explain women's subordination, its causes, and the ways in which it is perpetuated. It employs gender as a key concept, that is, the social relation between the sexes, arguing that this relation is not naturally given but socially constructed. Many feminists today think of gender as informing multiple dimensions of human experience, from individual identities to social and political institutions. Thus, gender operates through individual identifications, shared narratives, and symbols. As a social construct, gender not only produces a social reality but also signifies a relationship of power—the predominance of things and people gendered masculine over those gendered feminine. Thus, feminists also use gender as a methodological tool, a lens through which to look at the world in order to reveal gendered relationships of power. In international relations, feminist theory has made three distinct contributions: First, using gender as a lens, it has served to inform critiques of the field as gendered. Second, approaching gender as a social construct, it has described the reproduction of masculinity and femininity in the conduct of international relations. Finally, it has introduced a unique methodology that requires researchers to think about their own position in the research enterprise. This entry describes the three distinct contributions.

Compared with other disciplines, feminist theory entered the field of international relations relatively recently, with seminal writings by Cynthia Enloe and Jean Elshtain in the late 1980s and J. Ann Tickner, V. Spike Peterson, and Christine Sylvester in the early 1990s. While this first generation of feminist research in international relations focused on critiquing concepts, a second generation has focused on empirical investigations, which have resulted in a broader questioning of methodologies. International relations scholarship employing feminist theory has become institutionalized through the formation of the Feminist Theory and Gender Studies section of the International Studies Association in 1990 and the founding of a new journal in 1999, the *International Feminist Journal of Politics*.

Critiquing the Canon

Feminists have argued that standard concepts of international relations orthodoxy are gendered, that is, that they carry associations of masculinity and femininity. Classical realism imagines individuals seeking domination and engaged in a struggle for power. Feminists have suggested that this conception of individuals reflects masculine experiences. What realists consider human nature (following the theories of Thomas Hobbes) does not reflect the experiences of mothers and other caretakers who know themselves less as autonomous individuals separate from others but as connected in their caring relationships. Feminists point out that Thomas Hobbes's state of nature is populated by individuals who are self-generated, not born of or nurtured by women, but who—in his own words—shot out of the ground like mushrooms. This image has informed political theories that presume atomistic human subjects. Projected onto the state, it has become the starting point for a realist imagining of the international world as made up of autonomous and self-interested actors devoid of social connections.

In addition to constructing humans as essentially male, realist theories celebrate a militarized version of masculinity. They construct this form of masculinity as hegemonic, implicitly denigrating other forms of masculinity and any type of femininity. The opposition between hegemonic masculinity and subordinated masculinities and femininities drives narratives of war and security. For example, the opposition between a masculine protector and women in need of protection is a basic structuring device that justifies and thus perpetuates war making. In her examination of women and war, Elshtain gives further content to this opposition, finding that Western philosophy constructs the opposition as one between masculine "just warriors" and feminine "beautiful souls" imagined as innocent about the ways of the world. The opposition drives a basic narrative that has drawn women and men into war while silencing the stories of women and men that do not fit the stereotypes. The opposition also informs a multiplicity of stories around war fighting, from male courage and heroism to feminine nobility and sacrifice. By privileging an understanding of a militarized male as the prototypical human, theories of international relations produce an intrinsically biased, highly militarized understanding of the international world, while at the same time reproducing a thoroughly stereotypical understanding of what it means to be a woman or man.

In addition to defining gendered actors in international relations, international relations theory employs the gendered opposition between public and private in order to explain and prescribe particular types of conduct. For example, the separation of public morality from private disorder fuels Machiavelli's gendered conception of politics. *Virtú*, the character trait Machiavelli demanded from a leader, emerges as a masculine principle and is opposed to a feminized *fortuna*, that is, fate, which he thought needed to be conquered by male virtue. In political realism, *fortuna* reappears as the disorder of anarchy in the international system and the task of the policymaker is to tame this feminized disorder. A highly misogynist construction of gender relations thus cements the prescriptions of power politics.

If the citizen warrior reigns supreme in classical security studies, notions of rational man replace him in liberal international political economy (IPE). Feminists have criticized the instrumental rationality reflected in notions of economic man because it denies the neediness of human beings and their mutual dependence on each other. Presumptions of self-interested cost–benefit calculations in economic conduct neglect this dependence and underestimate the role of caring in the production of values.

As in security studies, feminists have also found explanations in IPE to be based on the denigration of the private sphere. This devaluing leads theories of political economy—in both the liberal and Marxist variety—to ignore the broad realm of social reproduction, the value generated, mostly by women, through the unpaid work of subsistence production, caring, and housekeeping.

Women and Men, Femininity and Masculinity in International Relations

The feminist critique of the international relations canon has directed scholars toward exploring areas left out of orthodox international relations scholarship. Arguably, feminist scholars operate within a different ontology of international relations—one that includes spaces and practices of women and one in which gender produces a diverse range of outcomes. A major focus of inquiries thus has been to make visible the practices of gendered agents. Another focus has been to analyze the situated constructions of femininity and masculinity that are produced in such practices.

In security studies, feminists have followed Enloe's lead and argued that power operates in a wide variety of everyday activities to sustain militarized security arrangements. Camouflage clothing, Rambo movies, the training programs of the U.S. military in schools and colleges, the military's regulation of prostitution around foreign army bases, the employment of sexualized and homely imagery when defense intellectuals talk about weapons systems—all function to construct the military and its often horrific weapons systems as a benign force, part of the normal order of things. Women gain a militarized role in this order as much as men—as mothers who encourage their sons to go to war, as pinup girls, nurses, prostitutes, and soldiers' wives who divert and sustain them and who remind them of the purpose of their fighting. Because war in this way depends on constructions of gender, the postwar reconstruction of countries needs to include a reconstruction of gender relations if it wants to prevent wars in the future.

Feminists also have explored the kinds of masculinities being constructed in particular wars and militaries, finding the warrior citizen model hegemonic but also tamed in the first Gulf War by a combination of toughness and tenderness. They have sought to destabilize the model of militarized masculinity by exploring the resistance against gays in the military and the military's refusal to acknowledge posttraumatic stress disorder. Denying differences in sexual orientation and denying the humanity of soldiers in the face of the horrors they encounter are efforts to preserve a militarized version of tough and fearless masculinity that thrives on the fiction of being the masculine protector of feminized "womenandchildren" (Enloe's formulation).

In IPE, feminist researchers have sought to make visible the work of women engaged in "reproduction," including subsistence farmers, maids, homeworkers, and sex workers. They have identified a different type of masculinity as hegemonic—that is, a bourgeois-rationalist model that celebrates competition, reason, and self-control. This model has become particularly emphasized in the context of neoliberal restructurings and globalization. Feminists have explored the construction of a class of global jet-setting managers empowered by forms of hegemonic masculinity and have shown that this has been accompanied by the development of a global low-pay feminized services economy. They have described the reconstructions of gender relations in postsocialist transition economies, including patterns that pushed women out of workforces and parliaments and re-enabled their sexual objectification. And they have interpreted the newly emerging East Asian developmentalism as a reversal of colonial hypermasculinity and as made possible by an inflated traditional paternalism in newly industrializing countries. Finally, they have explored feminist efforts to mainstream considerations of gender in institutions of global economic governance, such as the World Bank, showing how feminism becomes integrated into existing arrangements, sometimes improving the situation of women at the margins while also continuing to deny the relevance of the reproductive economy.

Feminist Methodology

Most feminists in international relations have argued for postpositivist methodologies that allow them to critically interrogate the field and its practices. Having grown out of a political movement, feminist research has often sought to ask questions relevant to women's experiences and produce knowledge useful to the movement. Rather than insisting on the researcher keeping a distance from her topic, feminists have sought to specify their relationship to the topic. For some, this requires defining a feminist standpoint—that is, specifying that they speak from the perspective of a specific movement or a specific community. For others, it has meant making into a subject matter their own position or experience during the research process. Yet others employ positivist methods that keep the researcher's distance from the object of inquiry but make sure to ask feminist questions. All share an element of reflexivity toward the research enterprise that distinguishes feminism from most other approaches in international relations. It allows feminists to problematize the political effects of the knowledge produced and arguably abide by a more robust version of objectivity than approaches that fail to keep in view the relationship between the researcher and the researched.

Elisabeth Prügl
Florida International University
Miami, Florida, United States

See also Feminism; Feminist Movements; International Political Economy; Positivism; Realism in International Relations

Further Readings

Elshtain, J. B. (1987). *Women and war*. New York: Basic Books.

Enloe, C. (1989). *Bananas, beaches, and bases: Making feminist sense of international politics*. Berkeley: University of California Press.

Peterson, V. S. (2003). *A critical rewriting of global political economy: Integrating reproductive, productive and virtual economies*. New York: Routledge.

Sylvester, C. (2002). *Feminist international relations: An unfinished journey*. Cambridge, UK: Cambridge University Press.

Tickner, J. A. (2001). *Gendering world politics*. New York: Columbia University Press.

Foreign Aid and Development

The Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) defines "development assistance" (or foreign aid) as financial flows that qualify as official development assistance (ODA). ODA is calculated as the sum of grants and loans to aid recipients that are (a) undertaken by the official sector of the donor country, (b) with promotion of economic development and welfare in recipient countries as the main objective, (c) at concessional financial terms, where the grant element is equal to at least 25%. In addition to financial flows, technical cooperation costs are included in ODA; but grants, loans, and credits for military purposes are excluded, and transfer payments to private individuals are in general not counted. The same goes for donations from the public, commercial loans, and foreign direct investment (FDI). It is common to treat ODA and foreign aid as equivalent, but this can be misleading. Assistance funded by nongovernmental organizations (NGOs) including church-based agencies has grown significantly in the past 25 years and now amounts to about one third of ODA.

Only official aid to "traditional" developing countries counted as ODA until 2005. For these (Part I) countries, there is a long-standing United Nations (UN) target from 1970 that they should receive 0.7% of donors' gross national income (GNI) as aid. Assistance to the "more advanced" Eastern European and "more advanced" developing (Part II) countries was recorded separately by DAC as "official aid," not included as part of ODA. DAC countries have over the years accounted for some 95% of all ODA flows, but the distinction between Part I and Part II countries is no longer used. All flows that fulfill the established criteria are now included in the aggregate measure of ODA, but NGO-funded contributions are not added.

In 2009, the total amount of ODA disbursed by donors to developing countries and multilateral organizations reached US$123.1 billion according to the OECD/DAC 2010 statistics. (For the data in this and subsequent paragraphs, see OECD, 2010).

This means that the average citizen in the donor countries contributed around US$149 as ODA. This can be compared with a figure of around US$64 in 1960–1973 and US$99 in 1992. However, the UN target of 0.7% of GNI is with few exceptions far from being reached. Donors disbursed 30.4% of total foreign aid to multilateral organizations in 2006, and some 70% of this flow was disbursed to developing countries, with the European Union and the International Monetary Fund and World Bank as the dominating sources followed by the UN and the Regional Development Banks.

It is a widespread perception that foreign aid amounts to a very significant resource, in both absolute and relative terms, and that aid is not insignificant measured relative to developing country production and income. At the same time, aid is much less sizeable when measured in relation to GNI or government budgets in the donor countries or in comparison with population size of aid-receiving countries. Moreover, aid has been on a declining trend since the early 1990s as a share of GNI in recipient countries. Most recipient countries receive aid to the order of 1.8% of their GNI per year with a median of 3.2%. This corresponds to a distribution of aid per capita with a mode of US$17.9 per year and a median of US$31. Accordingly, the relative size of the aid inflow varies significantly among recipients, and while the 13.2% size of the aid to GNI ratio in, for example, Tanzania may seem high, this share reflects not only the size of the aid flow but also the very low level of income. With this background, modest expectations are advisable when analyzing the overall impact of past aid on development.

In subsequent paragraphs, the historical and theoretical context of foreign aid is reviewed first. This is followed by overviews of the empirical evidence on the allocation of aid and its impact on furthering growth and development in aid-receiving countries. The conclusion provides discussion and summary remarks.

Historical and Theoretical Context

Foreign aid in its modern form has roots back to the early 1940s and intensified after the disruption that followed World War II. The international economic system had collapsed, and war-ravaged Europe faced a critical shortage of capital and an acute need for physical reconstruction. The response was the European Recovery Program, commonly known as the Marshall Plan. During the peak years, the United States transferred some 2% to 3% of its national income to help restore Europe. The Marshall Plan, which was administered by the Organisation for European Economic Co-operation (OEEC), the predecessor of the OECD, was implemented on schedule, and its success fuelled highly optimistic expectations about the future effectiveness of foreign aid.

After the success of the Marshall Plan, the attention of industrialized nations turned to the developing countries, many of which became independent around 1960. Economic growth in a state-led planning tradition became a key objective during the 1950s and 1960s, and it was widely believed that poverty and inequality would be quickly eliminated through growth and modernization ("trickle-down"). Theoretical thinking on economic development at the macrolevel was informed by both the Harrod-Domar growth model (which was extended into the two-gap model of Chenery and Strout) and the Lewis dual-sector model framework; whereas standard microeconomic cost–benefit analysis (CBA) was relied on in project analysis. The Harrod-Domar model is based on the assumption that investment is the key constraint on growth, whereas the Lewis model is concerned with the transfer of labor from the "traditional" (rural) to the "modern" (industrial) sector. CBA, on the other hand, assesses cost and benefits of individual microeconomic project interventions based on shadow prices for inputs and outputs and derives net present values and internal rates of return.

The major part of the rapidly increasing bilateral flows during the 1950s came from the United States. New bilateral donor agencies (other than the United States) were mainly established in the 1960s, with the Commonwealth-inspired Colombo Plan from 1950 being an exception. A transition toward more independent, multilateral relations also began to emerge during the 1960s. This created a constituency for foreign aid, and the non-aligned movement gave an articulated developing country focus to this voice, as did the various organs of the UN. The International Bank for Reconstruction and Development, or World Bank, established at the Bretton Woods Conference in 1944, came to play a central role in development assistance and international policy formulation, especially following the creation of the International Development Association in 1960.

The original Marshall Plan was built around support to finance general categories of imports and strengthen the balance of payments (i.e., program aid), but from the early 1950s, project aid became the dominating aid modality. Some donors continued to supply program aid, but aid was increasingly disbursed for the implementation of specific capital investment projects and associated technical assistance.

The multilateralism of aid became somewhat more pronounced after the mid-1970s. Multilateral channels were at the time seen as more efficient and less political than bilateral aid, so the UN, World Bank, and other multilateral agencies expanded their activities considerably. The 1970s also saw an increased focus on employment, income distribution, and poverty alleviation as essential objectives of development and foreign aid. The effectiveness of trickle-down was widely questioned, and new strategies referred to as basic human needs and redistribution with growth were formulated and propagated alongside more radical dependency theories of development. Nevertheless, the typical project aid modality remained largely unchanged.

The golden era of the 1960s and 1970s came to an abrupt end at the beginning of the 1980s. The second oil shock in 1979 reversed economic conditions, and there was a huge increase in interest rates due to the economic stabilization policies in the developed countries. The international debt crisis erupted, and macroeconomic imbalance became characteristic. On the political scene, Ronald Reagan and Margaret Thatcher came to power in the United States and the United Kingdom, respectively, and at the World Bank, Anne Krueger became vice president and chief economist, replacing Hollis Chenery. Economic circumstances in the developing countries and relations between the North and South changed radically. The crisis hit hard, especially in many African countries. Focus in development strategy and policy shifted to internal domestic policy failure, and achieving macroeconomic balance (externally and internally) became widely perceived as an essential prerequisite for renewed development.

"Rolling back the state" turned into a rallying call in the subsequent structural adjustment efforts, and reliance on market forces, outward orientation, and the role of the private sector, including NGOs, was emphasized by the World Bank and others. In parallel, poverty alleviation slipped out of view in mainstream agendas for economic reform but remained at the center of attention in more unorthodox thinking such as the "adjustment with a human face" approach of the United Nations Children's Fund (UNICEF). At the same time, bilateral donors and international agencies struggled with how to channel resources to the developing world. Quick-disbursing macroeconomic program assistance, such as balance of payments support and sector budget support (which were not tied to investment projects and which could be justified under the headings of stabilization and adjustment), appeared an ideal solution to the dilemma of maintaining the resource flow and the desire to promote policy reform. Financial program aid and adjustment loans (and eventually debt relief) became fashionable and policy conditionality more widespread. In other words, a rationale that corresponded well with the orthodox guidelines for good policy summarized by the "Washington Consensus" had been found for maintaining the aid flow.

Accordingly, total aid continued to grow steadily in real terms until the early 1990s, but after 1992, total aid flows started to decline in absolute terms until the turn of the millennium. Many reasons account for the fall in aggregate flows after 1992, including first of all the end of the Cold War. The same can be said for the weakening patron–client relationships among the developing countries and the former colonial powers. The traditional support of foreign aid by vocal interest groups in the industrial countries receded. Bilateral and multilateral aid institutions were subjected to criticism and characterized at times as blunt instruments of commercial interests in the industrial world or as self-interested, inefficient, rent-seeking bureaucracies. Moreover, acute awareness in donor countries of cases of bad governance, corruption, and "crony capitalism" led to skepticism about the credibility of governments receiving aid.

The potential role of foreign aid in all this attracted attention, and the fear that aid can generate undesirable dependency relationships became clear during the second part of the 1990s and persisted into the 21st century. In parallel, the perception that policy conditionality was failing to promote policy reform started to assert itself. This assessment prompted World Bank and independent academic researchers to start digging into the aid–growth relationship using modern econometric techniques. Even more recently, efforts to develop randomized program evaluation techniques (including a variety of experimental approaches) appeared based both on micro-CBA studies from the past and an increasing understanding of the need for developing proper counterfactuals in economic analysis of foreign aid and associated policy interventions.

Aid Allocation

Over the years, foreign aid has been justified in public policy pronouncements in widely differing ways, ranging from pure altruism to the shared benefits of economic development in poor countries and further on to the political ideology, foreign policy, and commercial interests of the donor country. Few would dispute that humanitarian sentiments have also motivated donors. Action following severe natural calamities, which continue to be endemic in poor countries, is an example. Food and emergency relief also remains an important form of aid. In addition, existing data and analyses confirm that donors allocate relatively more ODA to the poorest countries.

Emphasis on the needs of poor countries was a particularly prominent characteristic—and the underlying economic rationale—in much of the policy literature on foreign aid in the 1950s and 1960s. Here focus was on estimating aid requirements in the tradition of the two-gap model. While still influential in practice, the two-gap model has been subjected to a variety of criticisms, and in parallel, the role of aid has changed to a much more multidimensional set of concerns. Economic return is by no means the only goal of aid. Nevertheless, growth and economic development in aid-receiving countries have continued as a yardstick for the effectiveness of aid both in their own right and as necessary conditions for the realization of other development aims.

It is not new that selfish motives are critical in bilateral donor decisions. Moreover, bilateral donors do indeed behave very differently among themselves. Up to about 1990, the Cold War was used as a powerful justification for providing aid to developing countries to stem the spread of communism. Similarly, aid from socialist governments was motivated to promote socialist political and economic systems. Other strategic interests play a role as well. The United States has over the years earmarked substantial amounts of aid to Egypt and Israel; being a former colony is an important determinant in getting access to French aid; and voting behavior in the UN can affect aid allocation both bilaterally and through the multilateral system. The same goes for how bilateral donors are influenced in aid allocations by their own strategic and commercial interest versus the development motives of aid recipients.

It is widely accepted that not all donors behave the same; but the donor community has as a whole failed to meet the established international target of contributing 0.7% of their national income as ODA. This is so in spite of widespread endorsement of the recommendations for a large scaling up in the context of the Millennium Development Goals. Only the group of Scandinavian countries and the Netherlands (according to OECD/DAC statistics) have consistently met the 0.7% target since the mid-1970s, while the United States contributed around 0.2% of its GNI in 2009.

Aid Impact

To measure the effect of aid properly, the analyst must in principle be able to compare the value of a chosen indicator (such as growth or poverty reduction) in two strictly independent situations—with and without aid. To establish the "true" measure of aid impact, the importance of all other circumstances that have affected a given country over time needs to be properly accounted for. Comparing what actually happened with an appropriate counterfactual is the fundamental evaluation challenge. Yet there is in social science no way of addressing this problem (i.e., the challenge of establishing an appropriate counterfactual) in a broadly acceptable way without making assumptions that are bound to be debatable, in theory and in practice.

Accordingly, the past decades have witnessed a massive outpouring of studies on the effectiveness of foreign aid. This topic has been a central and recurring theme with which many development experts, subscribing to the different paradigms of development thinking, have grappled, and methodologies have varied. More specifically, (a) the impact of aid has been evaluated at the micro- and macroeconomic level, (b) cross-country comparisons as well as single-country case studies have been relied on, and (c) aid effectiveness research includes broad surveys of a qualitative and interdisciplinary nature as well as more hard-core quantitative work.

One key point on which there is at least some agreement in the literature is that aid has in many cases been highly successful at the microeconomic level. The most rigorous project evaluations in this area are done by the World Bank, and reports from the Independent Evaluation Group of the World Bank are generally encouraging. Average rates of return are generally above 20%, and decent project rates of return have been reported regularly over the years in one survey after the other. Overall, a mass of project-based evidence has been collected, and few dispute that aid interventions have worked in helping improve social outcomes through better health and in helping promote and develop appropriate technology (i.e., the green revolution) and so on. Yet it remains less clear what works in more concrete terms and what does not work, and doubts about aid's overall impact on growth and development linger on. The question is regularly raised whether all this adds up at the macrolevel.

It is easy to arrive at a negative association between aid and growth in simple aid–growth correlation analysis. There is, however, no logical inconsistency in development terms between little growth and aid inflows of the size experienced in the past. Donors allocate more aid to poorer countries, which are subject to difficulties and shocks of many kinds, including natural and man-made calamities. When countries have done well for a while average income has gone up, donors tend to transfer less aid, and eventually they withdraw. While such "graduation" may take a while, simple correlations are on this background likely to show a negative relationship; but they do not reveal the "true" impact of aid. Aid allocation matters, as do the major changes that have taken place in the global economy and affected the environment in which aid is implemented. Targets for aid have also been changing from one decade to the next. Thus, simple correlation analysis or storytelling cannot—and should not—be allowed to settle the causality debate about aid's potential impact on development.

It is, however, never straightforward to generalize from case studies, and this helps explain why macroeconomic cross-country (panel data) studies of the aid–growth link became so popular from around 1995. Such an approach makes it possible in principle to move beyond simplistic aid–growth correlation analysis, where the analysis of causal effects is rather primitive. Much of the modern empirical aid effectiveness literature has focused on whether the impact of aid is conditional on policy or whether aid can be expected to have a separate and positive impact, independent of policy. This has involved a mixture of concerns. They range from technically demanding econometric modeling issues to fundamentally different approaches to the design and implementation of development strategy and policy.

Overall, it has become clear that coming up with the "true" aid–growth relationship is far from easy, and aid is in any case of much too limited a size to turn the wheels of history. Yet, while aid is controversial, few reject aid as a potentially useful instrument in the fight against poverty. Careful, nuanced, and subtle assessments are advisable with the empirical evidence in hand; and the single most common result in the modern aid–growth literature is that aid seems to have had a positive impact on per capita growth. No excessive claims about aid impact on development should be made on this basis, and the empirical evidence remains an area of dispute, leaving the door open for conflicting interpretations and policy recommendations.

Conclusion

Foreign aid has been associated with development successes and failures, and the fundamental analytical problem in assessing its impact is that nobody has to date successfully identified the underlying development model. Analysts therefore continue to work with reduced-form models, which are bound to be debatable. In parallel, existing data suggest that foreign aid is far from equally effective everywhere. The necessary and sufficient conditions for aid to have a positive contribution on the development process remain elusive. In other words, how to come to grips better with what actually drives existing differences in the impact of foreign aid is a challenge in theory and practice. This is so, for example, in relation to potential interaction with economic policy, but the same goes for deeper structural characteristics.

The lack of generalized understanding of the complex links in particular country circumstances between aid, growth, and development objectives, such as poverty reduction, means that selectivity in aid allocation, based on simple macroeconomic criteria, is hard to defend. Few would argue that old-fashioned conditionality should be brought back to rule the way; but a better understanding of the intricacies of the donor–recipient relationship in theory and in practice would be valuable. This would as key elements include addressing issues such as (a) the best way to channel resources to the poor, when national governments are not capable and/or willing to take on this task; (b) how to ensure that aid delivered directly to national governments does not undermine local accountability; and (c) establishing the appropriate balance between aid going to the government vis-à-vis individuals and others in the private sector. Accordingly, how best to strengthen incentives in support of genuine domestic policy leadership is a challenge. The same goes for the fundamental task of furthering accountability and transparency vis-à-vis local populations.

Recent years have seen a drive to scale up aid. In this context, it is critically important to avoid making the mistake of the past of promising too much—that is, of contributing to the misconception that aid can on its own turn history. It would, based on history, appear that aid has much to offer, but managing expectations is far from easy. It is demanding to determine how best to make sure that promises made are kept. There are many unresolved issues here, including deciding how best to design incentives in aid agencies to meet this challenge alongside topics such as the role of independent evaluation, of coordination among multiple donors, and of the need to consider political economy issues, including the need to sharpen the incentives for recipients to use aid effectively in promoting development.

Finn Tarp
University of Copenhagen
Copenhagen, Denmark

See also Accountability; Bilateralism; Case Studies; Conditionality; Cost–Benefit Analysis; Hypothesis Testing; Interaction Effects; International Organizations; Modernization Theory; Neoliberalism; Nonstate Actors; Panel Data Analysis; Quantitative Methods, Basic Assumptions; World Bank

Further Readings

Arndt, C., Jones, S., & Tarp, F. (2010). *Aid, growth, and development: Have we come full circle?* (UNU-WIDER Working Paper No. 96). Helsinki, Finland: UNU-WIDER.

Burnside, C., & Dollar, D. (2000). Aid, policies, and growth. *American Economic Review, 90*, 847–868.

Cassen, R., & Associates. (1994). *Does aid work?* Oxford, UK: Clarendon Press.

Dalgaard, C.-J., Hansen, H., & Tarp, F. (2004). On the empirics of foreign aid and growth. *The Economic Journal, 114*, F191–F216.

Deaton, A. (2009). *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development* (NBER Working Paper Series No. 14690). Cambridge, MA: National Bureau of Economic Research.

Hansen, H., & Tarp, F. (2001). Aid and growth regressions. *Journal of Development Economics, 64*(2), 547–570.

Kanbur, R. (2006). The economics of international aid. In S. Christophe-Kolm & J. Mercier-Ythier (Eds.), *Handbook of the economics of giving, reciprocity and altruism* (pp. 1559–1588). Amsterdam: North-Holland.

Organisation for Economic Co-operation and Development. (2010). *DAC1 official and private flows*. Retrieved October 26, 2010, from http://stats.oecd.org/Index.aspx?DataSetCode=TABLE1

Organisation for Economic Co-operation and Development. (2010). *ODA by recipient*. Retrieved October 26, 2010, from http://stats.oecd.org/Index.aspx?DatasetCode=ODA_RECIP

Riddell, R. C. (2007). *Does foreign aid really work?* Oxford, UK: Oxford University Press.

Tarp, F. (Ed.). (2000). *Foreign aid and development: Lessons learnt and directions for the future*. London: Routledge.

Foreign Policy Analysis

Foreign Policy Analysis (upper case, abbreviated as FPA) is a subfield of the academic subject of international relations (IR) that has developed since the 1950s, broadly in parallel with IR itself. Foreign policy analysis (lower case, abbreviated as fpa) is what all commentators on matters of foreign policy do as a matter of routine, for one country or many. (This distinction between FPA and fpa will be used here for clarity.) Whether academics, journalists, or think tank researchers, they dissect the sources, goals, instruments, and feasibility of foreign policies and have thus gained over the years the sobriquet of *foreign policy analysts*. The term itself, however, is relatively modern in origin and owes much to the emergence of the academic field of FPA, which is the main focus of this entry, given its importance for IR and for political science more widely.

FPA has predominantly become established in the Anglo-Saxon world (i.e., England and her former colonies), including the United States, the United Kingdom, together with Scandinavia and some countries of the Commonwealth, such as Australia, Canada, India, Nigeria, and South Africa. Foreign policy has been long studied elsewhere but more in the fpa tradition and influenced more by history and international law than by political science. Even within the relatively narrow world of FPA, there are important differences of approach. The U.S. scholarly community has produced most of the important work in all areas of the subject, but it has also historically placed much more emphasis on large-scale data sets associated with the specialization known as comparative foreign policy (CFP), which has not attracted much interest (or the necessary resources) outside the United States. Elsewhere the focus has been more on case studies, on the microaspects of decision making, and on the ideas of foreign policy—particularly in the past decade, with the rise of constructivism and its concern with identity politics. But wherever the subject is studied, it is underpinned by the preoccupation with the relationship between process and outcome. Even those primarily concerned to generate events data have been motivated by the wish to discover correlations between certain kinds of foreign policy events and key variables, such as domestic upheaval or degrees of pluralism.

FPA can operate with a zoom or a wide-angle lens, on individual decisions, and on a class of actions—behavior in crises, for instance. Yet even close-textured analysis is set up in such a way as to enable comparisons to take place. In this, it is similar to the study of comparative politics, which can be regarded as its equivalent, or parallel, within mainstream political science. Yoked together, the two subjects constitute a solid bridge between political science and IR. Each also draws on history and area studies, as one cannot understand political culture, whether in the singular or the plural, without a grounding in its geographical and historical particularities. Theory is central to FPA, but it is not, by itself, a powerful tool in this particular context. FPA also looks to the study of international law and organization (more properly to the sociology of law and organization) for insights into the constraints under which foreign policymakers operate. In fact the literature of IR as a whole is useful in this respect. Even Kenneth Waltz's neorealism, which explicitly disavows having anything to say about foreign policy (it being a different level of analysis from his own), provides a clear picture of the nature of the international system, with implied guidance for the actors within it.

It might be thought that the comparison between FPA and comparative politics is misplaced because the former focuses on policy and the latter on the overall nature of a political system. But this is only superficially true. FPA has evolved beyond both policy and decision making to include all the domestic sources of foreign policy—and to some extent also the foreign sources of domestic policy. It thus provides scholars with a way of comparing how different polities behave outside their own borders and of analyzing the various ways in which domestic society becomes involved with IR. At one level, it is certainly the study of the external dimension of public policy and therefore directly comparable with that particular branch of political science. But it goes beyond that to engage with the nature of the political process as revealed through arguments about foreign policy.

If one looks for the intellectual origins of FPA, one needs to go back further than its actual beginnings, with the work of Richard Snyder and his colleagues at Princeton in the mid-1950s. Their work was rooted in the burgeoning belief in the value of social science, following in the powerful tracks of natural science, which had changed the world so dramatically over the previous century. This meant both that foreign policy should be analyzed in as dispassionate a manner as possible, avoiding the unproductive polemics that had marked the discussion of Munich, or McCarthyism, so as to probe beneath the surface of what is now called the "spin" from all sides, to reveal "what really happened." On this basis, governments and citizens could gain a better grasp of what was involved in improving the quality of what was from then on to be known as decision making—that is, a neutral term, divorced from the particularities of time or place, by comparison with public administration, which studied specific institutions such as Congress or Whitehall in their own terms. Given the dangers of the just commenced nuclear age this was an understandable, indeed noble, aim.

Beyond this general trend of the growth of social science, and its application to decisions about war and peace, lay also work in economics, politics, sociology, and philosophy going back to Aristotle, which had produced in various ways attempts to think systematically about how those in positions of political responsibility should—and did—calculate their best strategy under conditions of conflict, scarcity, and danger. Political theorists (as they are now called) had focused largely on how to promote the good life (or for whom) and how to manage power, often including speculation about relations with foreigners. Economists from the 18th century on had constructed theories of the efficient use of resources, while sociologists had begun to identify elites whose work extended beyond that of formal officeholders. The interest in psychology sparked by Sigmund Freud (together with the evident psychopathology of Adolf Hitler and his associates) was starting to turn a searchlight onto the behavior and motivations of all those in positions of power. Key individuals such as Max Weber were capable of making connections between these strands and thus laying down the interdisciplinary foundations necessary for the study of political decision making. It was therefore virtually inevitable that at some point the long tradition of concern over war and its consequences would fuse with wider developments in social science to produce a systematic approach to the study of foreign policy—in short, to produce *Foreign Policy Analysis*, even if the term itself took some time to become routine. In the following, the theoretical concerns, subfields, and contemporary significance of FPA are discussed.

Theoretical Concerns

As FPA has evolved, it has acquired certain defining characteristics, or key assumptions. These may not be shared by everyone working in the field to the same extent, but they are certainly points of reference that all must take into account. This section identifies five such assumptions and adds three more that may be in the process of attracting consensus as the field progresses and adapts to its changing intellectual and real-world environments. Between them, these eight assumptions delineate the central content of FPA while also helping mark its external boundaries.

The first, and most fundamental, assumption has already been referred to—namely, that *process affects outcomes*. This is why decision making is such an important part of FPA, even if it should never be confused with the field as a whole. The making of decisions and policies is studied because it is taken for granted that the way in which a problem is handled, and by whom, will have a considerable effect on the quality of the subsequent decisions. Perhaps surprisingly, before World War II, this point, which now seems a truism, was not explicitly formulated in such a way, even if much attention was given to the personality and foibles of individual leaders. But the structures in which they worked, to say nothing of their governmental cultures, was not given formal attention. FPA, however, soon made it clear (using among other tools, systems theory) that not only did decision making have certain common features across national contexts but also that feedback loops existed between policy outputs, reactions from outsiders, and a policy system. Thus, policy processes shape the outcomes in State A, but outcomes are also determined by interaction with States B, C, and so on, which in turn may modify the original structures for taking decisions in State A—and so on. One straightforward example is the setting up of the hotline between Washington and Moscow after the perilous missile crisis of October 1962—an innovation designed to improve the crisis management machinery in both capitals.

A second central characteristic of FPA is the acceptance that *foreign policy as an area is incapable of generating an overall theory of IR* in the way that work pitched at the level of the international system is. The latter may be deemed (especially by Foreign Policy Analysts) as parsimonious to the point of abstraction, telling us very little about the actual dilemmas of states and citizens beyond what can be deduced a priori from the overarching theory (as with neorealism), but at least it has a set of propositions about how the world as a whole functions. A focus on foreign policy cannot do this. Contrary to the view commonly expressed by Neorealists, foreign policy theories are not inherently reductionist, because they rarely claim to generate explanations of the international system. But they do focus on agency, that is, on how *agents*, or rather *actors* (the former being a more ambiguous term) cope with the various domestic and international structures in which they are embedded. This coping will be of a variable nature, according to the actor in question, even if it takes place within certain limits, dictated by what Edmund Burke called "the empire of circumstance." We may thus draw certain conclusions about what kind of behavior, whether by states or by other kinds of actor, meets with greater or lesser success, and therefore about what kind of world we all live in, but this is not the main purpose of FPA, and it is difficult to be rigorous about it unless we shift level to that of the system overall.

A related assumption to that just described is the consensus that has emerged over the past 2 decades in FPA that, even at the level of comparative foreign policy, it will not be possible to construct a convincing overarching theory—in other words, a single theory of foreign policy is the chimera. The reason for this is clear enough. Finding a single model of how foreign policy is made and conducted that would fit the 192 member states of the United Nations (to say nothing of nonstate actors) would entail such a high level of generality as to be banal in the extreme. When one allows for the importance of the vast range of different policy-making arrangements, degrees of interest group involvement, national cultures, and historical traditions, together with the mosaic of distinctive geopolitical locations, resource distribution, cultural patrimonies, and degrees of development of the states concerned, it is clear that an umbrella theory of foreign policy must be reduced to the same kind of lofty generalization as that imposed by realism or Marxism—indeed the latter can be justly accused of the reverse sin of reductionism, an error that might be labeled "holism" because of its flattening out of the variety of world politics in the interests of parsimony. That said, the flowcharts and taxonomies of the decision-making process that have been produced by such figures as Richard Snyder and Michael Brecher are extremely useful in revealing the wiring diagrams of what occurs in most foreign policy systems, to a greater or lesser extent—but they are not *theories*, that is, sets of propositions capable of explaining why certain patterns of behavior emerge and/or on giving guidance on how we should behave in any given situation.

The third point of reference for most FPA scholars follows from the skepticism over grand theory, namely the emphasis on *analysis*. But what is meant by this common term? It is less than theory and more than description, to be sure. It refers to breaking down the process of foreign policy making and execution into its component parts, to understand better what lies beneath the surface of events. To do this, it uses concepts and theories of a particular kind, as will be explained below, in a process of deconstruction and reconstruction. In part, this involves locating foreign policy in its own specific environments, identifying the constraints under which it thus operates, and also the interactions between action and environments that result. "Analysis" here thus has its routine, scientific associations, but it also has some of the connotations of psychoanalysis, in the sense of the wish to probe beneath the surface of events to uncover a semiconcealed, barely understood set of forces that shape behavior just as much as the formal roles of the actors or the reasons they give for their actions do. Most foreign policy analysts, whatever their own designation, present themselves as having this capacity to see beyond the public, or perhaps more exactly the public relations, dimension of international affairs.

The fourth underpinning assumption of FPA is that it generates *middle-range theories*. Middle-range theories are those that attempt to explain certain delimited and specified aspects of human behavior. They are self-evidently neither grand theory nor microtheory—the latter tending to attach to specific phenomena (or "phenotypes"), as with, for example, a theory that tried to explain Russian foreign policy in terms of the fear of encirclement. By contrast, middle-range theory addresses classes of phenomena ("genotypes") but within limits, so as to not overload the activity with too many variables. FPA has been notably successful at generating this kind of theory, because by focusing on, say, the operation of bureaucracy it inherently limits the scope of explanation to one set of players (bureaucrats) and to one aspect of politics (foreign policy). This allows for manageability without restricting intellectual ambition. While reality is complex, our understanding of it has to simplify, by theorizing the critical factors but not at such a high level of generality as to become meaningless—hence middle-range theory. There is no doubt, however, that this still leaves open the big issue of how to aggregate the various insights from diverse middle-range theories, just as we have to find some way of putting fruitfully together agent-based theories derived from FPA with structural theories pitched at the level of the international system. There is no easy answer to this dilemma. The most that can be said is that it is better to make sense of certain dimensions of reality and then to be faced with the problem of commensurability than to take refuge in excessive detail on the one hand or empirically thin generalization on the other.

The last of the conceptual anchor points for FPA has been implicit in much of what has gone before—namely, that the subject's principal focus is on *agency*, that is, the business of deciding and acting in the world. In one sense this seems a straightforward job description: Focus on what states and other actors do, rather than on the structures within which they operate. But the subject's own evolution, together with lessons learned from the philosophy of social science, has brought us to understand three important caveats. First, actors and structures are engaged in a continual process of mutual interaction and mutual constitution and therefore cannot be regarded as fixed points; second, some actors have rather more capacity to have an impact on structures than others do, just as not all structures are relevant to all actors in the same way; third, it is, indeed, not always possible to make clear distinctions between actors and structures. In part, this is because the latter are multiple and nest inside each other like Chinese boxes—for example, the French state exists within the structure of the European Union (EU) but is itself the main structure that French political parties relate to, while they in turn represent key structures for individual politicians and interest groups. But it is also because, while most actors are readily identifiable empirically, structures are often conceptual entities whose existence cannot be falsified, such as "the international community" or "globalization." Accordingly, it is more common for structures to be identified in terms of actors—whether dominant groups such as the United Nations Security Council, or "the West," in the case of the international community; or the aggregate of states, companies, and individuals, in the case of globalization. Waltz made a strenuous effort to avoid this trap in his neorealist theory by stressing that it was the *pattern* of the balance of power that determined the structure of international anarchy, even if this is all too easily reducible to the activities of major players.

The result of this problem so far as FPA is concerned is that it is all too easy to get drawn into wider issues than the problem of action and especially to neglect the viewpoint of the actor—when it is indispensable from the viewpoint of good scholarship to be able to reconstruct the perspective of decision makers themselves, rather than to impose on them assumptions drawn from a universal model. The other side of the coin is that foreign policy analysts may, in the very attempt to distinguish actors from structures, unnecessarily reify the state (or even actors like Al Qaeda), when they are more intelligently understood as a series of interconnected points of agency, overlapping with other apparent actors. This is clearly the case within the EU, where national foreign policy occupies a still important but nonetheless ambiguous position in relation to the Union's own collective activity. Ultimately, individual human beings make foreign policy decisions but always within organizational structures, some of which will have formal responsibilities and are accountable to various types of constituency. The nature of the agency that is the key interest of FPA therefore wobbles between and across the levels of individual and collective action, while the latter is itself divided between different political and bureaucratic formations and sometimes between a state and the alliances/organizations to which it belongs.

Beyond these five pillars of FPA research are three further assumptions that, while more contestable, are increasingly widely shared. The first of these is that *foreign policy is not only the business of states*. While it may seem something of a category mistake to talk of "foreign" policy in relation to actors that do not have sovereignty, or even a clear territorial expression, there can be little doubt that, among those described generically as nonstate actors (the use of the negative tells us of the difficulty in identifying any other common characteristics), many have political strategies for influencing IR. It is not unreasonable to conclude that they too possess international strategies that may be identified, with some qualification, as "private foreign policies." Certainly any organization, from the smallest firm to the United Nations, makes a distinction between its internal affairs and its external environment and has to formulate strategies for dealing with the latter. The British Academy, for example, which exists to promote scholarship, uses the term *Foreign Secretary* to describe the Fellow appointed to liaise with similar bodies in other countries.

The category of nonstate actors contains bodies as diverse as these civil society groupings at one end of the spectrum and Al Qaeda at the other. In between come transnational churches, corporations, sports organizations, and charities, many of which are taking on more of a political role, in terms of both the protection of their own interests internationally and the wish to influence the nature of what they perceive as a single, globalized, environment. Given that FPA is the main site of the concern with agency within IR, it must pay attention to the significance of these actors, their aims, instruments, and effectiveness. This is because of their inherent importance and also because of their interaction with states, shaping the latter's choices. Just as transnational relations has added another dimension to the traditional interstate focus of IR, so transnational actors have expanded the scope of FPA.

The proliferation of transnational actors has helped reinforce a further common preoccupation of FPA in recent years, that over *the increasing difficulty of defining foreign policy*. Apart from the fact that there is an inherent problem in deciding whether an aspect of public policy is by and large inwardly or outwardly directed, the development of the world economy since 1945, with the freer movement of goods, people, and capital, has meant that foreign and domestic policies increasingly merge into each other. Indeed, most ministries in developed states now work with their equivalents in other states, while classical foreign policy has to be concerned not simply with negotiation and the calibration of power but also (and quite legitimately) with the domestic politics of other states.

All this calls into question the boundaries of FPA and possibly its whole raison d'être. If the explanandum is only vaguely discernible, then how can a systematic analysis be possible, let alone an explanation? The answer to this is twofold: On the one hand, practitioners and citizens continue to use the term *foreign policy* as central to their political vocabulary. To put it crudely, if enough people think foreign policy exists, then it does exist or, at least, it is a sufficiently important idea to be worth serious investigation. On the other, whatever the blurring that has occurred between inside and outside, the two dimensions of actors' activity cannot be reduced to the same thing. Rather, and this proposition is likely to attract general support within the field, policy now exists on a single continuum, with matters that are almost entirely domestic (such as local government) at one end and those that are almost entirely foreign, such as a civil war in a third country where the given state has no particular interests, at the other. The task of FPA is then to interrogate areas of policy to see whether there are still important differences according to the extent of their external ramifications and also to show how international factors play into domestic politics (the "second image reversed"). In fact, the very enquiry into the nature and boundaries of foreign policy is an important scholarly task that many inside FPA have been engaged in.

The last of the more recent common reference points that underpin FPA derives from constructivism, and more from "thick" than from "thin" constructivism at that. This is the proposition that *identity* is central to our understanding of foreign policy, and (often) vice versa. Identity is an imprecise notion, even by comparison with those essentially contested concepts (e.g., national interest) that political scientists work with regularly. But its use does extend FPA's standard concern with the domestic sources of foreign policy to the area of culture, including nationalism, tradition, memory, and self-understanding. At first sight this simply adds another set of variables to the equation: now, if one wants a full understanding of a country's foreign relations, one has to take into account its sense of itself and its culture. Outsiders' perceptions of that identity will also be important. That this is hardly a new insight, indeed one that would have been familiar to Alexis de Tocqueville, matters little. The boost that constructivism has given to the study of culture and language has certainly affected FPA, both giving it a new lease of life and posing some new challenges to its orthodoxies.

The challenges seem largely to be at the level of methodology, in stressing discourse analysis, interviews, and texts more than foreign policy "events" or institutional data. More significantly, however, the "cultural turn" represents an epistemological challenge in the sense of the renewed criticism of positivism that it implies—not all FPA has been positivist in orientation, but it was undoubtedly a product of the "scientific" wave of IR work in the postwar period. It suggests that foreign policy can both arise from a constructed national identity and be constitutive of that identity. The circular tendency of this reasoning, not to mention its essentialism, is naturally not to the taste of everyone working in FPA. But a certain consensus is merging, as in IR more generally, to the effect not only that issues of identity and culture cannot be ignored but also that they may also be central to our understanding of much foreign policy behavior—not least in an era when religious strife has reemerged, and public diplomacy has become a more prominent instrument of state policy. Most would certainly concede that foreign policies such as that of Wilhelmine Germany, or Brezhnev's USSR, were profoundly important in determining not only the external image of those countries but their very conformation at the time.

The Subfields of Foreign Policy Analysis

Over its 50-year history, FPA has produced a number of well-researched subfields, sometimes drawing on work already done elsewhere in IR and political science but also often deriving from advances in the other social sciences. Taken together, they represent a massive advance on the understanding of foreign policy making that existed in the days of Edward Carr and Hans Morgenthau—which is not to say that the keen insights of figures such as these were misplaced; they were simply not cumulative, or part of a systematic schema. They also depended on the acceptance of a few top-heavy assumptions. Now, for anyone who cares to look, there is a vast array of analytic and comparative work on the machinery of foreign policy, on its conceptual aspects, on its implementation, and on its relationship to different kinds of polity. What is more, most of this work is accessible to intelligent general readers and is certainly capable of being translated into a form that enables it to enter political and popular debate. What follows outlines five important clusters of activity in the research agenda of FPA. No pretence at comprehensive coverage can be made, given the limited space available. But there cannot be much work of major interest that would not fit under one of the subheads that follow.

Leadership and Decision Making

FPA tends to start at the center and work outward. The center consists of the high officeholders, formally responsible for policy, and their intimate counselors. If we were living in the time of Louis XIV the focus would be on the king and his court. Now it is on the head of government, often but not always elected, together with his or her political advisers, the security council or its equivalent, and the heads of the key relevant departments, which means (at a minimum) the foreign and defense ministries, and the intelligence service. But what are the questions that FPA asks about this inner executive? The most straightforward is that familiar to political analysts over 3 millennia: What is the impact of an individual leader, extended in the post-Freudian age to include the impact of the leader's personality and psychological quirks, on foreign policy? An important line of thought that has emerged in FPA is the difficulty of considering such questions outside a relational context—that is, not only how does the leader behave and why, but how are his or her actions received by others? In short, this entails the analysis of leadership as such, which is central to political science as a whole, but has certain specific dimensions in IR, given that a good part of foreign policy making is conducted away from public scrutiny, while at the same time much of a leader's constituency actually consists of foreigners, whether other leaders or their publics.

Perhaps the most well-known product of FPA research relates to the psychology of small groups at the top level of the policy process—namely groupthink. This theory, which was validated by the various enquiries in the United States and Britain on the decision to invade Iraq in 2003, looks at the dynamics of leadership within the foreign policy executive, with particular reference to its psychological dimension and impact on the quality of decisions. It is closely related to our by now extensive understanding of how political perception works, with distortions caused by both emotional factors and formal roles within a group. FPA has learned much from its sister subject of political psychology but has also made a contribution of its own, notably in relation to groupthink, images of the foreign, and the concept of misperception, the consequences of which can be particularly dramatic in IR.

One interesting new strand of work that is emerging from this well-grounded tradition is that relating to neuroscience. It stresses the role of emotion and of preconscious factors in determining choice. Rationality is thus bounded not only by the limited ability of the human mind to process information or to trade off a large number of variables but also by emotional predisposition—some have said that we are hardwired, for example, to defend our existing beliefs. On the other hand, whatever the physiological and neurological sources of emotional preferences, it has been known from the early days of FPA that personal as well as political forms of bias affect decision making—indeed, that the two sources are difficult to distinguish. What the recent wave of interest in neuroscience has brought to the party is the insight that reason itself is intimately connected to emotion and that the two phenomena are not separated in the brain. The problem remains of how such distortions work at the collective level: Are they cumulative as in groupthink or reined in through the checks and balances of argument and interinstitutional conflict? This issue is at the heart of the subject as a whole and opens the possibility of advising practitioners on how best to avoid the wrong implications that may arise as a result of errors of such distortions.

It also connects up with the other main avenue of research on top-level decision making, that of bureaucratic politics. This came originally from practical insights under the Kennedy presidency, spawning a large and fruitful academic literature that has fed back into policy making in the form of direct advice but also as warnings to the citizenry that policy making is much more a matter of turf wars than they are led to believe. As a result, the insights of bureaucratic politics are now routinely used in case studies, in historical writing, and in think tank policy briefings. They do, nonetheless, cut both ways: On the one hand, they alert us to the way in which most outputs are the result of pulling and hauling, leading to arbitrary compromises, even in nondemocratic systems. This makes it virtually impossible to live up to the rational/realist ideal type of the consistent pursuit of a clearly identifiable national interest. On the other hand, they point up the advantages of collective debate, even on the crude basis of competing institutional interests. They can counter the personal biases of the leader and the tendencies to group conformity, which have been noted by the psychological approaches referred to above. This is both a matter of fact and one of preference, in that while some would wish to give a leader a free hand when dealing with foreign threats, others are convinced by the arguments for pluralism in decision making, in foreign policy as in other areas of public policy.

The Domestic Sources of Foreign Policy

One sign of progress in a social science is when its insights change from seeming marginal, or even outlandish, to being an accepted way of looking at a problem. This has been the case with FPA's emphasis on the domestic sources of foreign policy. Of course any claim has to be subject to the caveat of post hoc, ergo propter hoc, in that a change in general attitudes might have occurred anyway, through events, rather than because of any academic research. But FPA has certainly educated generations of students, against the still powerful realist skepticism over domestic factors, including the vagaries of public opinion, having the capacity to complicate, even shape, foreign policy. To a degree this has been a matter of life imitating art, as the more that an understanding of the foreign policy process has been disseminated, the more pressure groups have sprung up to insert themselves into the interstices of the process.

There have long been both hopes and fears about the way in which the public's opinion could disrupt or hold to account the best laid plans of their more expert leaders. On the pessimistic side were Alexis de Tocqueville and Walter Lippmann, while the optimists are liberals such as Woodrow Wilson or Senator Henry "Scoop" Jackson. Work done in FPA, however, has helped disentangle the analytical from the normative elements in this inherently political debate while distinguishing the new dimensions brought to elite–mass relations by the rise of the mass media, information technology, and the 24/7 news cycle. The activities of interest groups and social movements in relation to foreign policy have also been excavated, including the previously neglected gender dimension. Equally neglected for many years was the measurement of public attitudes on foreign policy questions, especially outside the United States. The creation of scientific polling by the Gallup Organization in 1937 led for the most part to the generation of data on elections, economics, and domestic political attitudes. When views were sought on international politics they were usually at a high level of generality or in the midst of major crises. But in recent years, it has become accepted both that the public can have a serious interest in foreign policy (with quite consistent views) and that the blurring of the line between domestic and foreign policy has ended the presumption that politics (and therefore public participation) stops at the national frontier.

Although Marxists and other political economists, together with some liberals, have always stressed the connection between domestic and international politics, within academic IR it has been primarily FPA that has insisted on its importance, not least in the face of the powerful neorealist movement of the 1980s and the rather different globalization paradigm of the 1990s, for both of which domestic events were subordinate to the logic of the international system. The constructivist wave has reinforced this research tradition by legitimizing a focus on culture and identity that had previously been seen as unscientific. Taken as a whole, FPA is a powerful demonstration of how realist attempts to abstract foreign policy from the political and social context in which it is embedded are misguided, even dangerously simplistic.

Organizational Logics

The theory of bureaucratic politics expounded by Graham Allison collapsed into one of his previously separate models of governmental politics and organizational behavior. There are, however, good reasons for keeping the two separate, as indeed the second edition of Allison's *Essence of Decision* (with Philip Zelikow, 1999) recognized. The institutional competition associated with the former has its most significant effects at the top levels of decision making. It also rests on the assumption that as individuals identify with their particular organizational home, they will pursue its particular interests or view of the world. By contrast, the organizational process approach identifies processes and mind-sets that derive from the nature of the complex state in the age of the masses. By contrast with the ideal type of Max Weber, the research conducted by and drawn on by FPA sees modern bureaucracy as developing pathologies that often inhibit the efficient and specialized divisions of labor it exists to promote. That is, the very conduct of government necessitates complex and often large organizations that inevitably fail to live up to their terms of reference. A century ago this would not have applied to foreign policy, with small and cohesive diplomatic services. Now it does, with many departments involved in policy making, defense budgets usually representing a major claim on the public purse, and the amount of information available about global politics growing exponentially.

The management of information, indeed, as of the increasingly large organizations themselves, is one of the major challenges facing the modern foreign policymaker and one on which FPA has much to offer. The study of intelligence is a case in point. For many years after 1945, intelligence was seen as both too sensitive and too difficult a subject for academic researchers. Over the past 2 to 3 decades, however, this assumption has changed, thanks to both historians and political scientists. The critical role played by secret intelligence—political, military, and economic—has come under scrutiny by all those interested in decision making. In part this has involved the study of turf disputes—for example, as between the numerous U.S. intelligence agencies—but it has also involved the more complex issue of how an issue gets framed and by whom, how to evaluate different sources of information, how to weigh capabilities against intentions, and—most critically in the case of Iraq in 2002 to 2003—what constitutes reliable professional judgment in a highly politicized environment. As the intelligence community has come to emerge from the shadows, it has become more willing to accept the value of dispassionate outside analysis of its role in foreign policy making.

Another benefit of the organizational approach comes via the focus on the decisional flow—namely, the stages through which major decisions pass. The orthodox approach in the political world is to focus on individuals, agencies, and the nature of public debate. But this needs complementing by the perspective to be found in FPA that peels back the surface layers to reveal the degree to which a particular decision, or policy, arose out of a linear process of information gathering and consultation or sprang from various forms of short-circuiting, whereby some parts of the government machine were cut out of the process and/or informal groups formed across organizational boundaries to hasten matters to the point of action. The linear process starts with information gathering or political inputs, moves to an assessment of the issue and the framing of the problem, which in turn lead to the delineation of options, the point of authoritative decision, and finally the complexities of implementation. Sometimes this model is actually borne out by events, but more often it represents an ideal type against which derogations can be measured. All too often the formal structure, on which democratic accountability depends (where possible in the first place), dissolves into a set of arrangements that is half constitutional and half ad hoc, with some stages of the process missed out or accelerated. Only careful analysis based on rigorous enquiry can reveal the true nature of the process. This will almost inevitably be after the event, with academics sometimes helped by the high-level enquiries often convened after the fiascoes that a high-handed approach to policy making can produce.

Goals, Purposes, and Values

In all agency-focused social science, the issue of the objectives of action is always at stake. What specific goals does an actor have, what broad purposes, and what underlying values? How far, indeed, is behavior purposive? We have seen that the study of both political psychology and organizational behavior places big question marks against any assumption of rationality, which may be extended to the idea of purposive behavior itself. Foreign policy is still a sufficiently discrete area for evidence on this question to be gathered systematically. Because goals for the most part have to be outwardly directed, against specific "others," they are more visible—or more notable when absent—than the objectives of domestic politics, which so often blur electoral considerations with those of economics and welfare. On the other hand, FPA has shown how foreign policy does not usually require legislation and can often fall back on maintaining a small number of parameters within acceptable tolerances.

The classic way of explaining the distinctive nature of foreign policy goals was through the concept of the national interest. It was one of the early achievements of FPA to blow holes in that concept. It has not sunk without trace, given the predilection of politicians for rhetoric and linguistic shorthands, but few serious observers would now regard it as unproblematic. In its place, the subject has emphasized the importance of understanding the ends–means relationship, the diverse and sometimes contradictory targets that decision makers pursue, and the partiality of many of the diverse interests that are pursued in the name of the state. The domestic dimension has been brought even more into play by showing how foreign policy relates to nation building and is infected by ideologies (including nationalism and religion) even in states that stress their pragmatism. In other words, the sources of both goal-directed and more impulsive behavior in foreign policy are regularly interrogated within FPA, both theoretically and by country specialists.

Closely related to this line of enquiry is the effectiveness of foreign policies, for success can hardly be measured without reference to the goals that are at stake. Research here has recently clustered around the concept of "soft power" ("civilian" or "normative" power in the case of the EU), but this is simply the latest stage in a long-running and cumulative program of enquiry. It derived from the realist preoccupation with exercise of power and steadily became more refined, variously seeing power as currency or on a spectrum between *Machtpolitik* and isolationism. It differentiated between the various instruments of foreign policy, with much attention being given to economic statecraft, to the use and threat of force, and to diplomacy in its various guises. Much of the most fruitful work has been done on the interplay and overlaps between these various categories, such as coercive diplomacy, sanctions, or the projection of a national image abroad (public diplomacy). Practitioners are particularly interested in academic findings on such subjects and frequently call on them during political debate.

For too long FPA had a tendency to concentrate on process and to avoid the issue of the content of foreign policy. It was forced out of this ostrich posture by the democratic peace debate, which had its origins more in philosophical approaches to IR than in the empirical study of state behavior. It did, however, quickly attract the empiricists and gave a further lease on life to the CFP school within FPA, whose large data sets on the behavior of classes of states seemed immediately relevant. The revival proved temporary, as once again CFP's combination of broad brush and methodological introversion proved insufficient to cope with the subtleties of the argument over what kind of democracies went to war in what kinds of circumstance. But there was a valuable payoff for FPA in that a new bridge was created between the empirical study of foreign policy and the renewed interest in ethics within IR. The work of the Carnegie Council and other bodies had long emphasized the need for a sophisticated discussion on the ethical dimension of international behavior. Now, following the lead of Stanley Hoffmann and Michael Walzer, this need began to be fulfilled, to the point where international theorists such as Chris Brown or Nicholas Wheeler regularly work on the edges of political theory and FPA, in their discussions of topics such as the motives for humanitarian interventionism or the priority to be given to overseas development aid. This has proved timely in that the post–Cold War era has led some governments to raise the ethical dimension of foreign policy without a trace of the defensiveness that would have been apparent in the years of Cold War realism and has liberated the forces in (usually Western) public opinion that wished to see foreign policy consistent with their own domestic values. Yet foreign policy analysts, as opposed to philosophers, bring an important practical perspective to bear on these complex problems. They relate moral questions to those of policy practice without falling back either on the simplistic realist assumption that everything ultimately reduces to the national interest or on the common liberal view that interests and ethics can be made to coincide. Rather, an analytical approach attempts to identify the multiple constituencies that a given state is trying to serve and to deconstruct the priorities that are revealed by practice. It can then go on to assess the feasibility of expecting X or Y of any given actor, given that actor's values and circumstances.

Environments and Structures

The external circumstances in which foreign policy is made constitute the last main area of work within FPA described here. From its early years the politically minded geographers Harold and Margaret Sprout had set the parameters for thinking about foreign policy within the frame of both material and political factors. In so doing they also raised the fundamental questions of causation and determinism—namely, how much freedom of action do foreign policymakers enjoy, given the difficulties of operating on the international level, and how much difference do size, location, and power make to the margins of maneuver. It might be thought, for example, that the great powers must enjoy more freedom than others, but much work within the context of FPA has been devoted to demonstrating the paradoxical freedoms of small states and the burdens of overextension on the part of large ones.

The Sprouts' emphasis was on the environments of foreign policy, and this has produced the useful orthodoxy of a division between the domestic, the external, and the psychological environments. They were also remarkably prescient in developing an ecological perspective on human affairs. To some extent, however, this perspective has been overtaken theoretically by the agent–structure debate. Structures are multiple and cross the internal/external frontier, but for foreign policymakers the most obvious structures in which they have to work are the international. These are material, in terms of the distribution of resources and climatic advantage; political but tangible, in terms of international law and institutions; and abstract/normative, in terms of the values of international society. The different ways in which they bear down on the great variety of kinds of states in the world are grist for the mill for FPA researchers, who can thereby work both from the outside inward and from the inside outward, showing how participation in the international system to some degree shapes the actors that operate within it and how the actors, both state and nonstate, singly and in groups, themselves determine the way in which the system evolves.

The Contemporary Significance of Foreign Policy Analysis

FPA as a field has three broad paths by which the problems outlined in this entry may be approached: theory, comparisons, and country studies. The theoretical activity revolves mainly around decision making, the problems of agency and rationality, the relationship between the state and the system in which it is embedded, and ethical questions. This is hardly a trivial agenda. In pursuing it, FPA borrows much from other subjects where advances have already been made but its own distinctive *problematique* requires original work that then makes its own contribution to social science—as we have seen with the now ubiquitous concepts of groupthink and bureaucratic politics.

The comparative perspective in FPA can operate on the grand scale, as with CFP, in an attempt to use all foreign policy events to identify patterns of behavior, and with correlations with types of state and/or types of internal events. More frequently, and more fruitfully, it now concentrates on limited classes of actor, such as small states, democracies, or specified types of events such as crises or economic sanctions. Even with such a restriction, the task of producing nonbanal generalizations is formidable and then most useful at a heuristic level. But despite the obstacles, all work in FPA is by definition comparative, if only at times implicitly, in that even single events are approached using common concepts and methodologies, so that they may be related to other similar instances. In fact, most case studies are influenced by the method of structured, focused comparison designed by Alexander George to minimize the randomness. Where processes are the focus rather than case studies the method is inherently comparative. The discussion of organizational process, or leadership, for example, begins with propositions of a universal character, which are then usually qualified according to the context in which they are applied. For example, is the domination of foreign policy by charismatic leaders, say, more likely in a developing country or in those with access to the sophisticated techniques of the mass media?

Country studies are the most conventional path for analysts to take. They are unfashionable because they are all too often seen as intellectually less interesting than theory or comparison, rather like the prejudice within social science against area studies. But this is a simplistic view. There are, after all, many ways of studying an individual country's foreign policy, some highly analytical and incisive. It is, in fact, rare these days to find a work that does not rise above the descriptive. Indeed, when significant change occurs in a state's international position, the natural response is a flood of work within the broad FPA tradition—as was the case with Germany and the United States after 1989 and China after its rise to economic power. This tradition is indispensable, for however insightful a theory, or however rigorous and wide-ranging a work of comparison, they cannot provide the rich, in-depth understanding of a particular country and its foreign policy choices. Any attempt to deduce the strategy of Burma or Saudi Arabia, for example, from neorealism or some other parsimonious theory, is doomed to superficiality. Such states are comparable only up to a very limited point, qua states. Individual cases might fit a generalization in some broad sense, but the resulting knowledge cannot have the thickness achieved by combining theoretically based analysis with expertise in the country or particular issue at stake.

This is why FPA, like all other aspects of social science, cannot be a subject unto itself. It relies on collaboration with others. If IR and political science are the most immediate points of contact, other social sciences, notably psychology, geography, and political economy, are equally important. What is more, FPA needs to be close to history, philosophy, and area studies, which may not be social sciences themselves but provide much foundational material for the latter. But FPA is not just a matter of scope, a field where the master disciplines play out their games. Over half a century, it has focused attention on one of the most important areas of human activity, if judged by the potential consequences for the mass of humanity. And it has accumulated a considerable body of knowledge, and techniques, for throwing light on foreign policy wherever it may be practiced. For that reason, it is important in itself, as a source of intellectual progress, and for the help it can give all of us, practitioners and citizens, in the search to understand the marvelous but often dangerous world we live in.

Christopher Hill
University of Cambridge
Cambridge, United Kingdom

See also Diplomacy; Foreign Aid and Development; Governance, Global; International Relations, Theory; International Solidarity; Risk and Public Policy

Further Readings

Allison, G., & Zelikow, P. (1999). *Essence of decision: Explaining the Cuban missile crisis* (2nd ed.). New York: Longman.

Baldwin, D. A. (1985). *Economic statecraft.* Princeton, NJ: Princeton University Press.

Carlsnaes, W. (2002). Foreign policy. In W. Carlsnaes, T. Risse, & B. A. Simmons (Eds.), *Handbook of international relations* (pp. 331–349). London: Sage.

George, A., & Simons, W. E. (Eds.). (1994). *The limits of coercive diplomacy* (2nd ed.). Boulder, CO: Westview.

Hill, C. (2003). *The changing politics of foreign policy*. Basingstoke, UK: Palgrave Macmillan.

Hudson, V. M. (2007). *Foreign policy analysis: Classic and contemporary theory.* Lanham, MD: Rowman & Littlefield.

Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes* (2nd ed.). Boston: Houghton Mifflin.

Jervis, R. (1976). *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.

Khong, Y. F. (1992). *Analogies at war: Korea, Munich, Dien Bien Phu and the Vietnam decisions of 1965*. Princeton, NJ: Princeton University Press.

Nye, J. S. (2008). *The powers to lead.* Oxford, UK: Oxford University Press.

Rosenau, J. N. (1971). *The scientific study of foreign policy.* New York: Free Press.

Smith, S., Hadfield, A., & Dunne, T. (Eds.). (2008). *Foreign policy: Theories, actors, cases.* Oxford, UK: Oxford University Press.

Snyder, R. C., Bruck, H. W., & Sapin, B. (Eds.). (2002). *Foreign policy decision-making revisited.* New York: Palgrave Macmillan.

Wheeler, N. J. (2000). *Saving strangers: Humanitarian intervention in international society*. Oxford, UK: Oxford University Press.

Wolfers, A. (1962). *Discord and collaboration: Essays on international politics*. Baltimore: Johns Hopkins University Press.

Formal Theory

See Game Theory; Social Choice Theory

Freedom

See Liberty

Functionalism

Functionalism refers to the methodological concept of explaining social phenomena by specifying an asymmetrical relationship between two objects under consideration. Functionalism rests on the view that phenomena can be best explained in terms of what they do and what their impact is on other phenomena; thus, it considers systems of interaction among individuals and groups. In political science, functionalism emphasizes the functions of social institutions. After providing a general definition of functional explanation, this entry outlines the ideas of some of the most influential originators of functionalism. Next, it describes the main stages in the development—structural functionalism, equivalence functionalism, and neo-functionalism. The entry ends with a critical summary of the status of functionalism today.

Definition

A functional explanation can take this general form:

> Given a system S in a certain state s with a structure T, there is an activity a from the point of view of the observer, regularly coming from an element E of T, and having an effect on S or its environment.

The theoretical status of *s* can either be stationary (the most frequent case) or dynamic. If one can, with respect to a certain theoretical point of reference, claim a systematic relationship between *a* and *s*, then *E* can be interpreted as a "function" for, or *a* as a "functional contribution" to, the maintenance of *S*, whether it means stability, identity, equilibrium, or changing of *S*. This explanation of *s* neither involves an explication of *a*'s origin nor the causal nexus of its effect on *s*. For example, a political party can be considered as contributing to the working democracy without constructing this function as the cause of its creation.

Functional explanations are based on empirical evidence: After *s* and *a* have been observed, the functional character of the relationship between them can be deduced from the features of *T* and *E*. Of course, these explanations can be formulated with a teleological orientation: as a projection of a future *s*, derived from *T* as a goal-orientated agent, under the assumption that *a*'s effect on *S* is a functional prerequisite for the maintenance of it. In case there is a deliberately role-differentiated organization under consideration, such a prognosis will become normative. It will express the expectation that *S*, as a whole, will function in the way prescribed by its blueprint.

However, functional explanations make sense only under several conditions. For example, if one states that general and free elections in parliamentary systems have the "function" of maintaining the circulation of political elites, then a bottom-up nomination of candidates is assumed, but other functions (such as lawmaking or representation) are not, however, excluded. Functional analyses are based on theoretical assumptions, but they hardly suggest new ones by themselves. According to the widespread usage of functional explanations, systems can be represented as institutions, religious behavior patterns, cultures, societies, and so on, whereas *a* may operate as a kind of activity, exchange, information, sanction, service, coercion, production, and other forms of output.

Origins

Functionalism, as a tradition or "school," culminated in the 1950s and 1960s. However, its origins can be found in the works of many of the classic social and political theorists. Charles de Montesquieu may be considered such an early proto-functionalist thinker. In his chief work *Spirit of Laws* (1748), he classified the components of political systems in terms of their physical, geographic, climatic, mental, and cultural "nature"—that is, the structure of their external systems. The respective legal principles activate the "natural" structures toward an outcome of political order by which the subjects are reminded of their duties and rights. The privileges of nobility, for instance, are a function of its freedom, whereas parliamentary power is one of the principles of constitutional monarchy. At the same time, Montesquieu tried to make it clear that important political functions can often hide behind ostensibly incidental or useless symbols and ways of acting.

An important predecessor of classical functionalism was Herbert Spencer, who, as a leading social Darwinist, tried to explain the development of industrial society in an evolutionist way. According to Spencer, the law underlying societal evolution proceeds from homogeneity of equal and independent parts to heterogeneity of unequal and dependent ones. These parts—institutions, groups, technologies, ideas, and so on—tend to become increasingly specialized and, as a result, fulfill increasingly special functions for society. The starting point of this process is that the management of population growth implies prerequisites such as advances of productivity, distribution, or regulation. Social differentiation, functional specialization, and interdependence of parts feed back to one another as mechanisms of the same movement.

One of the founders of modern sociology, Émile Durkheim, then adopted functional analysis, logically distinguishing it from genetic-causal explanation, as a methodological program. In his works *The Division of Labor in Society* (1892) and *Rules of the Sociological Method* (1895), he argued:

1. In explaining a social phenomenon, one has to distinguish its generating cause from the function actually fulfilled by it. It is inadmissible to explain collective phenomena in a utilitarian way because their respective functions can originate from or serve different purposes. At the same time, teleological-causal explanation is not promising because a collective phenomenon such as society does not constitute a consistently acting whole.

2. With this, Durkheim suggested that we understand social phenomena as emergents: They come into being through an accumulation of immense courses of aggregated individual actions or of unintended effects of purposeful action.

3. A social phenomenon emerges and tends to enhance the importance of its function when individuals recognize the advantage gained from it and develop an interest in optimizing profit. Durkheim was one of the first to conceptualize function as a self-reinforcing mechanism and a feedback system. Durkheim demonstrated this in the light of division of labor. Specialization of occupations and social roles lead actors into growing dependence. With this, human networks become closer, making the exchange of specialized goods easier. As a result, a social order that Durkheim called *organic solidarity* is established, as a kind of moral sui generis. In the course of this process, people also experience the secondary benefits of division of labor, such as the growth of economic productivity or differentiation of the legal system. Finally, modern division of labor derives from needs that have emerged at a preceding state. Thus, the theoretical meaning of social progress is a growing complexity of functions.

In the 1930s through the 1950s, functionalist thinking spread, especially in anthropology. From an ethnological point of view in the Western perspective, there was a strong temptation to interpret "strange" phenomena by taking the stability of the social and cultural systems as a point of reference. Bronislaw Malinowski (1944) and Alfred Radcliffe-Brown (1952), for instance, tried to explain the performance of magic rites by Pacific Islanders in this way. While Malinowski emphasized the physio-psychological functions of need satisfaction and fear reduction by magic, Radcliffe-Brown claimed that magic served to promote the survival of the group and the maintenance of the system's structure. It was George Homans who presented an interesting proposition for "reconciling" these approaches by overcoming the one-sidedness of both of them: Magic initially was performed to soothe ghosts or to release stress and anxiety in dangerous situations, such as sea fishing or childbearing, and to bring about confidence, provided that the rites are performed in the "right" way. At the same time, this procedure involves conformity with the norms that are in force, the violation of which would have serious consequences for the group. Therefore, society instills in its members a fear of punishment for noncompliance through rites that serve as a means of enforcing compliance with such norms. At the same time, magic rites are important group activities, strengthening the bond and solidarity of the group and promoting the survival of the system thereby. Homans, by combining the psychological and the structural aspects of a functional explanation, demonstrated how to take the hybrid nature of functions into account—following Durkheim's basic idea.

Structural Functionalism

Classical functionalism culminated in Talcott Parsons's (1951) systems theory. Parsons called his approach *structural functionalism*, later on giving up this term at the zenith of his work. Social systems, as the point of reference for his theory, have to solve a limited number of general problems for their maintenance or keeping their balance. When modern society is being considered, "function" refers to the solutions to such general problems. Some functions are externally directed: They provide for the adaptation of the system to its environment. Others are internally directed, prerequisites for the integration of parts or for actors' motivations. In his famous AGIL (adaptation, goal attainment, integration, latency) scheme, Parsons systematically specified the four main system problems and related them to the four subsystems, from which the corresponding functions of problem solution proceed (see Table 1). The functions are interrelated by processes of interchanging special classes of "media" that define their respective outputs, as if they were material, psychic, and social goods or energies in the metabolism of society.

Along these lines, adaptation—that is, maintenance within environment—is the function of economy, including technology, labor, and consumption, and the output of it flows in the general form of (equivalences of) money. Attainment of goals fundamental to society is guaranteed by the subsystem of polity bringing about effects of control in the form of power. The interchange between A and G, for instance, can be manifested in the way laws and sanctions (polity) regulate the functioning of markets, whereas the economy enables

Table 1 Talcott Parsons's AGIL Model

Function for Solution of System Problems	*Subsystem*	*Media of Interchange*
(A) Adaptation	Economy	Money
(G) Goal attainment	Polity	Power
(I) Integration	Societal community	Moral appeals
(L) Latency (pattern maintenance)	Fiduciary system	Influence

Note: AGIL = adaptation, goal attainment, integration, latency.

the functioning of the political system by providing revenue from taxes. Societal communities contribute to system integration by providing for an optimal climate of morals, solidarity, and socialization as well as sanctioning conformity with values and norms (e.g., by awarding prestige). Finally, pattern maintenance typically arises from the fiduciary subsystem, the actors of which are sources of influence toward ensuring common interests, standards of professionalism, and other grounds of trust. Within the subsystems, social roles—the intersections of structure and actor—are interlinking. They represent the institutionalized expectations toward the occupants of social positions derived from the given structure of social differentiation. Therefore, the classification of functional parts according to the four system problems is repeated on every sublevel: Subsystems consist in functionally structured subsystems of the same nature. The system as a whole is vaulted by the cultural system as the total of legitimate values, norms, and symbols.

For several decades, Parsons's functionalist approach, as a heuristic scheme, has been used in various research applications. However, to apply the AGIL model to empirical systems involved some compromises. For example, some overlapping of the subsystem–function dimensions had to be accepted in matching a specific function with a given part of the system.

Most criticism of Parsons's approach focused on its inherent tendency of holism to consider society as an "actor" without any microfoundations, assuming the fiction of a system's tendency toward an equilibrium, the fiction of common goals of society, and their teleological character. Finally, Parsons was said to have explained social change too simply by referring to interference from environment or internal accident, instead of taking into consideration dynamic features in the system's structure itself.

Robert K. Merton was the most prominent critic of structural functionalism. Not rejecting functionalism in total, he suggested revising it on many points. First of all, he distinguished between "manifest" and "latent" functions. The former arise as cumulated effects from individual, intended action, mostly a case of individual conformity to collective rationality. The latter means the type of collective effects that emerge beyond the control of individuals, when there is no matching of goals and consequences. This distinction corresponds to the difference between "functional systems" and "systems of interdependence" suggested by Raymond Boudon (1981). The first are goal-orientated systems of interaction regulated on the base of reciprocal, specialized role expectations as well as controlling by means of input–output assessments. Business organizations, universities, political parties, bureaucracies, and so on are of this kind. On the other hand, systems of interdependence emerge wherever actors are interrelated, losing (parts of) control over their actions, ultimately becoming involved in a complex structure that is nontransparent both in its nature and in its consequences. If, for instance, at an election with three contesting parties, a clear ranking of votes is the result, one usually would expect that the leader of the most successful Party A is likely to become the next prime minister. Yet in case of a proportional representation, together with a proved culture of forming coalitions, the next prime minister could come from the secondary Party B as well if Party B successfully negotiates with Party C, which has the least number of votes, and together Party B's and Party C's votes total more than 50%. The main feature of this system of interdependence is that cabinet formation does not derive directly from the ranking of votes. Thus, the ideal function of democratic elections to provide for a freely selected government is not achieved. To understand this structure of emergence, it is necessary to find out the statistical as well as institutional and procedural rules that are interdependently responsible for the composition of collective consequences out of the individual effects of action.

Merton had already approached such a skepticism of the fruitfulness of trying to functionally analyze systems of interdependence, including society. He turned away from "big theory" (as a global explanation), pleading for progress in developing less abstract, and less general, "theories of the middle range," which target a more modest explanation. As a consequence, he proposed deconstructing the general concept of function and making a distinction between function (in a sense of adaptation or adjustment of a system), dysfunction (as a consequence of action destabilizing a system), and nonfunctional consequences (irrelevant to the state of a system). For a functional analysis with higher concreteness and specification, he proposed a number of conditions or restrictions such as the following:

- the phenomena under consideration should have an institutional character and be standardized;
- subjective moments of action should be taken into account as well as the distinction of positive, negative, and neutral consequences;
- the functional affected units should be specified as well as the criteria for system maintenance;
- functional alternatives should be specified, with respect to their scope of application;
- dysfunctions producing change should be identified;
- comparative studies to better validate approaches of explanation should be performed;
- features of actors, meanings, places, motivational, and overt behavior that are involved in pattern of action should be specified.

It is hard to imagine that this demanding program would have been executed in a more or less exhaustive way. The usual restrictions on personal, material, and temporal resources and on theoretical potentialities allowed only for selective implementations in a few studies.

Equivalence Functionalism

Along with claiming to have established a universal theory of social systems as modern, Niklas Luhmann (1995) proposed a new definition of "function" that allowed for a larger scope of interpretations. In his approach, functions have both a multiple sense, as with Durkheim, and also an "equivalent" one: Functions serve as criteria for comparison of the actually identified functions with virtual ones—that is, some of the theoretically possible relationships in a system under consideration. They are viewed as an expression of "unity and difference," indicating contingent structures of the system. Luhmann used the research question of what mechanisms regulate the shortness of goods to illustrate his approach. At first sight, it is easy to come to a function as a combination of moral rules and economic mechanisms, the particular effects of which can be compared with each other. Yet the question of why shortness actually shall be regulated "can only be answered with respect to system/environment differences" (Luhmann, 1995, p. 404, German edition). The openness, indefiniteness, and complexity of this concept of function involve a comparative approach for functional explanations, distinguishing particular system levels as well as system–environment differences.

Neo-Functionalism

Jeffrey Alexander (1998) and others made the next step in the further development of functionalism, calling their approach *neo-functionalism*. The main aspects of their revisions, in accord with some of Merton's suggestions, are as follows:

- Parsons's analytic model shall be preserved, however, there were no explications given on the methodological status of the concept "function" or a functional explanation.
- Parsons's view of society as a network of relatively independent parts tending toward an equilibrium is not empirically substantiated.
- The inflexibility of the systems theory approach must be loosened up by the inclusion of "contingencies" in individual action: The interrelations between cultural order, as the frame of action and the contingent moments in themselves, are to be empirically analyzed. Thereby, knowledge from micropolitics or sociology, ethnomethodology, symbolic interactionism, and so on must become integrated.

- Parsons's "idealistic" view, that is, the predomination of culture and values in modeling social order, is widened by emphasizing the instrumental–rational dimensions of action and by incorporating conflict theories.

Alexander gave an example of analyzing social change based on these criteria. The general acceptance of new institutional structures that, on the basis of approved principles, are better able to resolve an acute problem will emerge at the end of a multistage, conflictual process. As a result of the pressure from this process, competing, contrasting interest groups will form and be stimulated to action. Then it becomes functional to optimize an internal division of labor so the numbers of alternatives are reduced step by step, and a generalization of public discussion emerges. If, farther on, under the pressure of "institutional entrepreneurs," it happens that an established pattern becomes illegitimate, then a certain alternative can win if it is accepted as final by the actors. At the same time, the general value system was never jeopardized.

Recent Variants in Political Science

Variants of (neo-)functionalism are still taking effect in parts of comparative politics. Here, the concept of function mostly refers to the meaning of institutional task, concern, responsibility, or another aspect of purposeful action carried out by or expected from actors, systems, or strategies. Thus, functions are understood as intentional rather than emergent processes or collective effects of action. Furthermore, approaches in political functionalism are based on empirical generalizations rather than on explicit theoretical or methodological principles. Common to them is the conception that political systems or processes, such as international affairs or European unification, can be controlled better by frontier-crossing institutions, nongovernmental organizations, experts, or supraregional networks of scientists than by governments within their territorial sovereignty. Processes of political decision chiefly run along the actual activities (functions) of these agents instead of those done by formally legitimated authorities, who only validate the results. Therefore, certain needs or interests of groups, who are the truly operating agents, have become the focal point of research, so far making use of a bottom-up method of analysis.

If the functional output of an institution oversteps a certain threshold value, then a spillover effect may emerge, for instance, when the outcome of international economic cooperation involves an enhanced political cooperation (A. J. R. Groom & Paul Taylor, 1975; David Mitrany, 1975).

A reversing of functional thinking can be found in the theoretical framework for comparative politics by Gabriel Almond, Bingham Powell, Russell Dalton, and Kaare Strøm (2008). After having subjected a great number of states to comparative analysis, a limited typology of functions, taken as inherent processes of system maintenance, is generated and standardized for all of the cases. Then the same procedure is done with respect to the system units exercising functions, such as parliaments, executives, courts, bureaucracies, and so on. Due to their origin and direction, the functions are distinguished in the following way. *System functions* are settled in the domestic environment: socialization, recruitment, and communication. They are intervening conditions for the *process functions*: interest articulation and aggregation, policy making, and policy implementation and adjustment. *Policy functions*, as a result, operate as the essential outputs: extraction, regulation, and distribution. They again react to the process functions, thus closing a dynamic control loop, which, beyond that, is interchanged with the political systems of other states. In empirical practice, after having operationalized these concepts for the empirical application, comparative analyses are carried out to find out how, in different real systems, the functions under consideration are fulfilled and what this means with respect to policy production as well as to the kinds of relationships between the states under observation. To state it once more: As a feature of this approach, functions do not mean ex post facto interpretations of relations having been observed. Rather, they are outcomes from institutions and agents, a priori specified as indicators for data collection, that on the ground of foreknowledge are available to the researchers.

Conclusion

Functionalism as a scientific school or continuously developing research program has had its day. The main arguments against it were too serious to permit further promising development. Among the problems were the following:

- an inclination to holism, with a neglect of effects of action on the micro- and mesolevel;
- an overemphasis on system stability or equilibrium, without explicating potentials of change as a feature of structure itself;
- a temptation to look at systems by analogy with organisms—that is, institutions may function in principle like the organs of a living being;
- a tendency to make teleological assumptions, even with respect to systems whose structures are not created purposefully; overabstractedness of functions having been discovered, for lack of specification of their origins, places, and ranges of manifestation; lack of specification of the actors involved; or lack of efficacy; and
- a lack of comparative studies for the purpose of better validation of function detected as well as the identification of functional equivalences.

Nevertheless, many suggestions from classical functionalism entered the mainstream of social and political science, such as the basic distinction between genetic-causal and functional explanations, focusing one's attention to undesirable collective consequences of individual actions, finally sharpening the view to a complexity and interaction that defy simple interpretations.

Hartmut Lüdtke
Philipps-Universität Marburg
Marburg, Germany

See also Durkheim, Émile; Interdependence; Methodology; Political Sociology as a Field of Study; Systems Theory

Further Readings

Alexander, J. (1998). *Neofunctionalism and after*. Oxford, UK: Blackwell.

Almond, G. A., Powell, G. B., Jr., Dalton, R., & Strøm, K. (2008). *Comparative politics today* (5th ed.). New York: Pearson Longman.

Boudon, R. (1981). *The logic of social action: An introduction to sociological analysis*. London: Routledge & Kegan Paul.

Durkheim, É. (1984). *The division of labor in society*. New York: Free Press.

Groom, A. J. R., & Taylor, P. (Eds.). (1975). *Functionalism: Theory and practice in international relations*. London: University of London Press.

Luhmann, N. (1995). *Social systems*. Palo Alto, CA: Stanford University Press.

Malinowski, B. (1944). *A scientific theory of culture, and other essays*. Chapel Hill: University of North Carolina Press.

Merton, R. K. (1968). *Social theory and social structure* (Enlarged ed.). New York; London: Free Press/Collier-Macmillan.

Mitrany, D. (1975). *The functional theory of politics*. London: M. Robertson.

Parsons, T. (1951). *The social system*. London: Routledge.

Radcliffe-Brown, A. R. (1952). *Structure and function in primitive society*. London: Cohen & West.

FUNDAMENTALISM

Fundamentalism is an inclusive term for modern religious movements that appeal to past events, texts, and authoritative figures and project into the future a variety of doctrines, stories, or laws that protect the group that is devoted to them and that serve to motivate and influence attitudes and actions, more or less aggressive, toward those outside the group.

The term may not be acceptable to adherents in all religions and cultures because it was first identified with 20th-century American Protestantism. However, the concept may be described as portable and is easily translated to cognate movements in the various religions and cultures. Sometimes it is appropriate to speak of these as "fundamentalist-like" movements, or movements that bear "family resemblances" to fundamentalism.

Historical Background

The term *fundamentalism* did not appear in dictionaries or encyclopedias before the 20th century. Scholars trace it to usage by partisans in denominational conflicts in the United States, debates that divided especially the Baptist and Presbyterian groups. Thus, among the Northern Baptists, an editor complained that great numbers in the church body claimed to be and wanted to be called "conservative," but their conservatism was passive. As such, it was not useful in the decisive struggles of the period. He and those who sided with him described a situation in which there were theological and moral threats to the religious identity and integrity of the adherents.

In the case of the American Protestants, the threats came from some "modern" or "modernist" claims and innovations. These threats, which included the challenge posed by Darwinism to the inherited biblical understandings of the creation of the universe and of humans, were seen to be an assault on human dignity; to undercut faith in the authority of the scriptures, which were described as inerrant in all detail; and to erode the psychological and theological boundaries of the members of the group. The leaders who rallied those who would, in the terms of the day, "do battle for the Lord" named their cause "fundamentalism." When the term, which they bannered in the beginning, invited stigmatization by their churchly foes and in the public media and consciousness, some of them chose other terms to signal that they were less aggressive. In Protestantism, the term *evangelical* was sometimes used.

From the beginning, it was clear that "fundamentalism" was not the same as "conservatism," though both may have started from the same base. Fundamentalism was a specific response to modern challenges; the word *reaction* best serves to describe the impulse and strategy of fundamentalists. The modern assaults on this reaction were corrosive and erosive of boundaries around the group and took the form of increased openness to and tolerance of other beliefs on the part of adherents.

It could be said that all through history religious groups have felt threats to their beliefs, practices, and identities, so one must ask why the early-20th-century challenges met the particular forms of response that they did. While it is never possible to substantiate exact claims regarding cause-and-effect relations, some elements stand out. One of the most notable is the sophistication of mass media of communication, notably the radio. Such media made it possible to reach beyond immediate locales to recruit, minister to, and draw on the discontents and aspirations of the like-minded who lived far away. Such employment of media suggests one of the paradoxes of fundamentalisms: their use of the most modern instruments to propagate messages that they advertise as and may believe to be what some Americans called "the old-time religion."

In the context of other religions, there has been an identification of fundamentalism with anticolonial and nationalist programs. Thus, in India, around 1925, the Rashtriya Svayamesevak Sangh (RSS) was formed to link religious nationalist movements with mainly Hindu bases; this was extended by the creation of the Bharatiya Janata Sangh (now the Bharatiya Janata Party, BJP) in 1951. In Egypt in 1928, the Muslim Brotherhood organized to resist what its leaders saw as secularizing forces and was promoted as a fusion of the Islamic religion and nationalist goals. These leaders revived classic originating stories from the Muslim past, building on the Koran, the revelatory text uttered by Allah through the prophet Mohammed. Those who followed the founders of the Muslim Brotherhood set out to restore, enhance, and enforce laws from the sharia, the code of Islamic laws. From Egypt, Wahabi Muslim movements fanned out, for instance, to the Arabian Peninsula, where they found ready followers.

Characteristics of Fundamentalisms

Scholars who view fundamentalisms synoptically and comparatively find characteristics common to most of them. When such scholars point to similar motives and expressions in, for example, American Protestantism and Shiite Islam, they are not saying that fundamentalisms are all the same. To suggest that would be foolish, because all fundamentalisms must have something to be fundamentalist about. The myths that animate Hindu movements differ vastly from those of Judaism or Islam. These myths are usually autochthonous, which means that they grew up out of their own soil and were not transplanted.

Devotion to Scriptures

The host documents in most fundamentalisms are ancient scriptures, almost always believed to be a revelation of God or the gods, as in Hinduism, and thus carrying divine authority. When some religious scholars began to interpret scriptures as they would other ancient writings, they appeared to some believers to be the ultimate threats, and their questioning and assaults inspired more precise defining defense than was needed when there was no challenge. When various schools of interpretation emerge, often competitively, it is necessary for the movements to make claims for one uniquely. Since these texts have been cherished and guarded by the community that gave rise to them, they also come to be used as definers of exclusivity. All other sacred or putatively sacred texts are repudiated. No measure or bases for tolerance or synthesis, attitudes typically born in the Enlightenment era, is permitted.

Group Exclusivism

If scripturalism is the first mark, then the exclusivism just mentioned is a necessary correlate. Psychologists who study fundamentalisms note that, across the board, fundamentalisms draw people who have no tolerance for ambiguity, contradiction, or paradox. If those who have nothing at stake in the question of their authenticity or who are from another religion find such phenomena, fundamentalist scholars immediately offer explanations that appear to define away any apparent contradictions. The drama of interpretation that excites many in religion is closed down in the interest of a monopolistic, unitary interpretation within a community, where only one version of one story from one text is seen to be "fundamental" and inerrant.

Rejection of Moderates

These texts may have vastly different characteristics. In Judaism and Islam, fundamentalisms are legalistic; an adherent lives by the divinely revealed law. In Christianity, in part retaining its Jewish base, communities thrive on stories and legalistic application of norms inspired by the texts. When critics challenge the integrity of a sacred text, they cause distress within the groups that revere the text in question. Couple this with intolerance of ambiguity, and it can be seen that the texts impart rules and meanings that cannot be questioned. For this reason, in many fundamentalisms, moderates are more despised than modernists. The modernist is an obvious enemy. The moderate knows the norms of the group but is open to listening to other sources and, in a sense, is thus traitorous, which means giving the game away. The apostate, the questioner, and the person at the margin demand the most defensive response and inspire positive counteraction.

"The Fundamentals" as Weapons

Such action includes a closing of the ranks to keep those enclosed within and to protect against influences from without. Fundamentalisms can be seen as enclaves—self-reinforcing and bounded communities. Those committed to each other around the fundamentals need spiritual ammunition to use against the enemy, the "other." This includes key doctrines, stories, or descriptions that are important to the believing community and a negation of all others. To further the chosen causes, fundamentalisms need and have rationales for setbacks and postponements of various consummations, such as military victory, the conversion of multitudes to the faith, or the assurance that God is on one's side.

Confidence in Victory

However the goals of the community are defined, no matter how great the odds against success may be, and despite setbacks and even the martyrdom of some members, the group has the assurance of ultimate victory: God or the forces behind the texts and the movement may not always be visible or clear, but they cannot withdraw the promises extended to the group any more than they could withhold the truth in the texts and usually the ancient laws and meanings.

Conversion of Others

If fundamentalists define "the other" negatively and are ready for defensive and aggressive action, sometimes lethal in character, they do not ordinarily allow their exclusivism to be absolute; there is an appeal to others to whom members witness and who come to see the assets that are enjoyed by those within the group, an appeal that leads to calls for conversion. The concept of being "born again," which is one of the marks of Protestant fundamentalism, is enriched and enhanced as members spread the message, utter threats to the future of the resistant, and offer those who join rewards that may be realized in this life or in a life to come.

The Politics of Fundamentalisms

Most fundamentalisms expect an eschatological consummation; that is, their adherents await an end to history and the world as it is now known and, more immediately, an end to their present earthly existence. For Jews, the ancient expectation of a Messiah—the anointed one who would rescue Israel and give it an exalted place—gets translated in the contemporary world into political actions that justify and impel actions in the modern state of Israel. The ultratraditionalist Haredim, orthodox Jews who are extreme in their attempts to fulfill the law, do not support the state of Israel as a fulfillment of Messianic dreams and are not considered by friends, foes, or scholars to be fundamentalists, to whom they, nevertheless, bear some likeness. Some version of rewards or fulfillment in the life to come, usually offered to the community of Israel, is another version of this eschatological hope.

Islamic fundamentalism adheres fervently to calls to serve Allah in ways that will ensure a future to Muslims as individuals and within the community. One's personal death, particularly if it occurs in such service, is followed by a fulfillment in the life to come. In fundamentalisms, that service acquires political and often military involvement, so that the values and meanings of the Islamic community are realized.

Christian fundamentalists may hold different views of the future beyond history, but they do have distinctive views. Most of these are related to the prophecy of a millennium, a 1,000-year reign of Christ, who will return to the earth. Christian fundamentalists also share with other Christians a belief in a personal resurrection in the life to come and proclaim one or another of the biblical pictures of what such a realized life involves.

Accenting both personal and communal consummation beyond history, as nonfundamentalists in these faiths also may do, led and leads many to be classified as "otherworldly." That is, the true drama of their existence transcended an ordinary this-worldly life. The main task in proclaiming the truth of the religion was to convert others, which meant to rescue them from the travails of earthly life. In this stage of their development, such fundamentalisms do not seem to have much to do with politics. Earlier histories of their movements tended to pick up on their otherworldly emphasis and show them to be unconcerned with the drama of politics. A typical case was that of a founder of American Protestant fundamentalism, Dwight Moody, who announced his program thus: The world, he said, was a flood. God gave him a boat and told him, "Save all you can." That was it.

Of course, that was not wholly it, in Protestant or other fundamentalisms. Most of these movements spent much energy in inculcating personal virtues that can serve as benefits in the *polis*, the human city, and the political order. All things being equal, one would observe them to be law-abiding and conforming to approved norms of society. Most of them have the reputation for being rigorous, in Christian terms "puritanical" and in Muslim versions more so. Fundamentalisms stipulate rules of behavior and expect and enforce adherence to them. They engage in the critique of violations of what they take to be the norms of good civil society. Where the regimes within which they live have what they consider to be positive characteristics, they are supportive of civil goals. Most of them are patriotic and quite characteristically can be militant in support of their state. Few fundamentalists are pacifist, and they believe that their God placed the sword in the hands of the state, so it is to be used against offenders from other tribes and states and nations. God is a warrior God, justifying the actions of God's people when their ways are congruent with divine commands.

If it were possible to quantify types within observed fundamentalist groups, it is safe to say that most of them concentrate on individual vices and virtues just as they focus on otherworldly realizations. Yet there are enormous and dramatic exceptions, and late in the 20th century, most were no longer perceived as simply otherworldly. How the turn came about differs in various religions and nations, but it is informative to concentrate on vivid examples. Thus, American Protestant fundamentalists, living generally peaceful ethical lives on the sidelines, rather suddenly changed and, in alliance with conservative evangelicals, Roman Catholics, and Jews, formed "moral majorities" or coalitions that advanced very pointed political programs and garnered support for them. One notable Protestant fundamentalist, Jerry Falwell, expressed a rationale for the change. Years before he turned political, his camp criticized Christians who were political and thus were sinful, since politics was tainting and distracting. *Now*, however, for the churches *not* to be in politics was the sin.

Why the change? In the political order and cultural ethos, threats had come against which private fundamentalists had no defense; the mass media in particular were universally intrusive. Fundamentalists responded to these cultural changes by urging those who shared their beliefs to work together to change that "corrupt" order through politics. Fundamentalists produced some political leaders or gravitated toward them, exploiting them just as they were often being exploited. Fundamentalists were not unique among religious groups in this respect. What was remarkable was that the believers who had first and foremost opposed political involvement now mastered many techniques, forms of organization, and patterns of rhetoric to make gains for their causes and themselves.

This turn to the political occurred in part in response to specific public actions. To follow the American Protestant example, the threat of communism during the Cold War called them to militancy. U.S. Supreme Court rulings on the legality of abortion and the illegality of state-sponsored prayer in public schools did much to solidify and motivate their base and attract converts. Their political action was generally compatible with the goals and means of other religious groups in a republic: One garnered votes, promoted legislation, sought constitutional change where necessary, and "won some and lost some," as is the manner in politics, which involves compromise and partial victories or defeats.

Where the terms of governmental life are not those of republics, fundamentalists have had to resort to other kinds of tactics. Most visible in the final 2 decades of the 20th century and the beginning of the 21st were the stirrings in Islam. Concurrently, Hindu-based fundamentalist movements in India engaged in military and terrorist action against Muslims, and vice versa. Still, Islam produced movements of such size and scope compared with other fundamentalisms and attracted so much media attention around the world that "Islamic fundamentalism" came to be the best known—probably the most feared by its enemies—and the most inspiring of all to its constituencies. Many of these fundamentalists were theocratic, like the fundamentalist Jews, who promoted settlement in the West Bank to fulfill their vision of a Biblically based map of the state of Israel: They believed that God ruled and had revealed that divine purposes were to be fulfilled through their actions, directed against "the infidel" or the "heretic."

The goal of some Muslim fundamentalists was to realize a dream they had nurtured for centuries when Muslims were not rulers or where they had once ruled but now had been pushed back. What one scholar, Marshall Hodgson, called "Islamdom"—on the model of "Christendom," when and where Christians had a domain (*-dom*)—was something to be retrieved from the infidels, who had conquered parts of it or corrupted all of it. The Iranian Revolution under Ayatollah Khomeini in 1979 to 1980 was of this character: The United States must be attacked when and because this "Great Satan" supported the enemies of those who would rule wherever Muslims had once ruled.

In radical Islamism, such terrorism is justified as following in the light of divine commands and promises, mediated through texts rich in story, or doctrine, or law. Terrorism is thus an extreme form of action beyond politics, since politics always involves some compromise, and fundamentalisms are absolutist and uncompromising. That is the characteristic of fundamentalism that is unsettling to nonfundamentalists in government or to those who inhabit other nations, hold on to other religions and customs, and would like to find ways to live with people of other religions.

Not all fundamentalisms can be or need be pictured as static and unchangeable. Some movements among them move on, thanks to the company they keep in coalitions, the realism that sets in when goals are frustrated, or their devotion to texts other than the militant or terrorist texts in Hebrew Scriptures, the Koran, the Christian New Testament, or a variety of Hindu texts. Using them, they would preach and promise peace and understanding.

Martin E. Marty
University of Chicago
Chicago, Illinois, United States

See also Islam; Judaism; Theocracy

Further Readings

Almond, G. W., Appleby, R. S., & Sivan, E. (2003). *Strong religion: The rise of fundamentalisms around the world*. Chicago: University of Chicago Press.

Jurgensmeyer, M. (1993). *The new cold war: Religious nationalism confronts the secular state*. Berkeley: University of California Press.

Lawrence, B. (1989). *Defenders of God: The fundamentalist revolt against the modern age*. San Francisco: Harper & Row.

Marty, M., & Appleby, R. S. (1993). *Fundamentalisms and the state: Remaking politics, economics, and militance*. Chicago: University of Chicago Press.

Marty, M., & Appleby, R. S. (1995). *Fundamentalisms and society*. Chicago: University of Chicago Press.

Marsden, G. (1980). *Fundamentalism and American culture: The shaping of twentieth-century evangelicalism*. New York: Oxford University Press.

Saha, S. C., & Carr, T. K. (2001). *Religious fundamentalism in developing countries*. Westport, CT: Greenwood.

Fundamentalist Movements, Islamic

The term *fundamentalism* has its origin in a struggle that developed between conservative and liberal Protestant Christian theologians in the United States around the turn of the 20th century. This struggle was epitomized in the publication by the Bible Institute of Los Angeles in the years between 1910 and 1915 of the series *The Fundamentals: A Testimony to Truth*, which in 12 volumes detailed 90 fundamental beliefs on which there could be no compromise for a true Christian believer. Later, the term took on a broader, more general meaning of strict and uncompromising adherence to and advocacy of certain basic principles or a set of beliefs, though it is most commonly used in reference to religious movements that are presumably characterized by such an attitude and that serve to motivate and effect aggressive attitudes and actions toward those outside the group. These include evangelical movements in the United States and elsewhere, fundamentalist Jewish settlers on the West Bank, and similar contemporary developments. In the international public sphere, and not the least in the media, *fundamentalism* is currently very often used to describe political movements among Muslim populations that legitimize their ideology and politics through a reference to Islam. Other terms in use for this phenomenon are "political Islam" and "Islamism," both of which have the analytical advantage that they carry a somewhat lesser baggage of preconceptions. The following, therefore, mostly focuses on recent religious and political developments in the Islamic sphere.

What is involved in the political utilization of the mobilizing resources of the Muslim culture and religion? In understanding the phenomenon, it is important to distinguish clearly between two levels of analysis: on the one hand, the reasons why political actors have recourse to framing their ideological and political discourse with reference to semantically potent symbols from Muslim culture, history, and religion, sometimes to the extent of insisting on using a distinctly Islamic vocabulary, and, on the other, the extremely varied uses to which these same actors put those Islamic references in national and international politics and the variables that affect their choices.

Reasons for the Use of Islam as a Reference for Guidance and Meaning

At the first level, there are two central driving forces behind the recourse to Islam as a reference for guidance and meaning: One is related to the defense of Muslim identity against a perceived foreign onslaught; the other is linked to an attempt to resurrect a moral framework for a rapidly changing society.

Maintaining Muslim Identity

In the context of the colonial expansion in the 19th and 20th centuries, the symbolic categories and signs of the imposed colonial culture came to replace the institutional and normative, and even the aesthetic, expression of the local Muslim culture. The "losing" culture found itself progressively marginalized and "indigenized" or "folklorized." The exclusion of its symbolic attributes from the production of meaning and more specifically from expressing values perceived as universal served dramatically to underline the humiliating centrality of the colonizing culture. While confining the expressions of religious culture to the sphere of personal status, the importation of secularization also initiated a very different process in "the lands of Islam." In the European context in general, secularism often appeared as a liberation from religious power and as part of a modernization that facilitated the coexistence of potentially conflicting loyalties and identities. In the Muslim world, where secularism was imported by the colonizing power or by elites that were associated with it, it has carried the sense of a denial of the local culture and its right to satisfy the normative needs of the society in the direction of public affairs. This privilege has been brutally transferred to a foreign culture, not the least through a wholesale importation of judicial codes.

It is against this background of a symbolic dispossession resulting from the hegemonic drive of the culture of the colonizers that a generation of Muslims in the period between the two World Wars committed themselves to restoring the visibility of the inherited Muslim culture and reaffirming the centrality of its norms. It seemed to this generation that modernization should no more be constructed around the opposition between "progress" and "reaction" or "dominant class" and "popular masses" but between "endogenous" and "exogenous" and "inherited" and "imported" cultural traits. With an indigenous vocabulary, the Islamist rhetoric brings to the dynamic of nationalist resistance to foreign domination a new symbolic resource. In rehabilitating the use of references that had for a time been excluded from political discourse, this rhetoric at the same time allows for the reconciliation of this discourse with a more intuitive culture. This culture is above all perceived of as that of the ancestors (real or mythical), with whom a symbolic affiliation now again becomes possible. The restoration of the categories of the Muslim culture does not produce more than a relative rupture. What it does is reaffirm the ambition of the endogenous culture of adding its contribution to the expression of universality on an equal footing with other cultures of the world and in particular with that of the former colonizers.

At the start, in the face of what a part of society perceived as a wave of rampant Westernization, the Islamist mobilization consisted of reaffirming the political role of the religious endogenous culture in the resistance to colonization. In 1928, 10 years after the dismembering of the Ottoman Empire, 4 years after the dissolution of the Caliphate, and 8 years before the recognition of a precarious independence for Egypt (Treaty of London, 1936), the founding in Ismailiyya at the Suez Canal of the Muslim Brotherhood by the schoolteacher Hasan al-Banna may be seen as the beginning of the organization of this reaction. In the context of the anticolonial struggle, the Muslim Brothers devoted themselves to reaffirming the centrality of the religious lexicon in the fight for independence. In Egypt, they engaged simultaneously the internal political debate and the field of social and educational work. Adept at modern modes of mobilization such as public meetings and the publication of magazines, they also very rapidly developed regional ambitions. In 1948, the Brothers in Egypt mobilized volunteers to counter the creation of Israel, and this became the beginning of a rapid spread of resistance throughout the countries of the Arab Middle East. This first Islamist generation, however, failed to achieve control of the state apparatuses that were eventually left vacant by the colonizers. While some old royal dynasties managed to reassert themselves in many central countries of the Middle East, the reins of power were grasped by nationalists of a more secular leaning, typically through military coups.

From the period when most countries in the Middle East gained independence (1945–1970) until the 1990s, two overlapping processes unfolded: first, the affirmation and then the beginning of the challenging of the indigenous elites that had come to power. The primary critique the Islamists directed toward the Nasserists and Baathists in power was that they suffered from a deficit of cultural authenticity. The Islamists called for the continuation in the ideological and symbolic sphere of the distancing from the colonizers that had until then been focused on achieving independence in political and economic terms. In the Maghreb, the persistence of the official and public use of the French language made this rupture explicit: The secular elites progressively became denounced as *Hizb Fransa*, "the party of France."

Except on the Arab peninsula and in Jordan, the first generation of Islamist mobilization was eventually not only denied access to the legal political scene but was systematically suppressed. Even more than at the time of its reaction to colonial violence, it was faced with suppression when the first radical ideological and political offshoots saw the light and won their first followers. The crushing of the Muslim Brothers by Nasser in Egypt in 1954, after he had realized the strength of their mobilizing potential, is the archetype of this repression. The "theology of war" that Sayyid Qutb—imprisoned continuously from 1954 until he was hanged in 1966—developed in response to Nasser's torture is central to the revolutionary posture that from the 1970s onward was adopted by radical groups. To respond to the injustice of the tyrant who oppressed him, Qutb declared him an infidel, a *kafir*, and thus deprived him of the protection afforded to members of the Muslim community. He further denounced the tyrant as the standard bearer of a new *jahiliyya* (a term usually referring to the state of ignorance that prevailed in Arabia before Muhammad). Then, he expanded the scope of his excommunication and did the same to the society that refused to reject such a ruler. His rhetoric was picked up in diverse forms by those who some years later (October 6, 1981) assassinated the Egyptian president Anwar al-Sadat, and to this day, it continues to mobilize the devotees of "direct action" simultaneously against the infidel rulers and against their foreign protectors. The Muslim Brotherhood also joined the opposition forces, which led to the downfall of President Mubarak in 2011. The emergence of the radical Qutb-inspired groups formed part of a much wider phenomenon: the great wave of Islamic awakening that swept the Muslim world from the mid-1970s onward and that greatly reinvigorated the Islamist movements. The vastly expanded numbers of high school and university students were the main basis of the early stages of this awakening. Some of these students were attracted to the violent groups, but most eventually contributed to swelling again the ranks of the Muslim Brothers, who were by now staunchly settled on a course of peaceful work for political and social change, except in the fight against a foreign occupier, as in Palestine. As for the radicalization that would lead a fringe of the wider movement to commit acts such as the September 11, 2001, attacks, the ideological variable is most often invoked in explaining it, and emphasis is put on the diffusion of the sectarian theology of Qutb among the activists trained in Afghanistan. But this inverses the causal order and obscures the centrality of simple political variables. The radical drive is a result of the unilateralism of the existing political order. First, there is that of the Arab regimes. For a whole generation of opposition members, the promises of the 1980s of a democratic transition have brutally given way to a repressive backlash. Despite some hesitant opening of the institutional political sphere in some countries for a while, the generalization of the use of torture at the national level and of the recourse to "hard power" by the regional (Israeli) and global (U.S.) powers worked to discredit reformist and legalist strategies for change and thus reinforced the credibility of the sectarian currents and their revolutionary strategies.

From 1990, Islamist parties here and there found doors being tentatively and partially opened to the legal political arena. Some (in Jordan, Algeria, Palestine, Lebanon, Kuwait, Sudan, and Morocco) established more or less permanent footholds in the parliaments—occupying ministerial posts (Yemen) or even directing governments (Palestine). But in the global arena as well as in the terrain of the Israeli–Palestinian conflict and in most national frameworks, these openings turned out to be limited indeed. Almost everywhere, elections proved the fragility of popular support for the regimes and incited them (in Morocco, Algeria, Tunisia, and Egypt in particular) to narrow the limited openings that had been granted and to define political competition as a threat to security.

The prevailing unilateralism also operated at the global scale, where the militarization of the foreign policy of the United States echoed the regional generalization of repression and torture. At the Sharm al-Shaykh summit in March 1996, under cover of the "fight against terrorism," the cooperation between the U.S., European, and Israeli security apparatuses on one side and the repressive apparatuses of the Arab regimes on the other reinforced the ideological criminalization of all resistance to the existing local and global order.

Providing a Moral Framework

But Islamism is a multifaceted movement rooted in the legitimate concerns of broad layers of the population. A sort of identity politics mobilized to counter foreign domination and repressive regimes allied with the foreigners is a central driving force.

Yet the success of Islamism in dominating the political and ideological landscape of popular activism is also linked to another strong dynamic at work, which may be termed the promotion of modernization under pious discipline. Looking at the leading activists of the movements in question, one finds that they are representative of dynamic elements within the middle and lower-middle classes, for whom modernization opens the way to upward social mobility. They have a positive attitude toward economic and technological modernization. But the ensuing social change is rapidly breaking up the old framework of town quarter, village, kinship group, and patron–client relations in which people led their lives. To counter the uncertainty and fear of social chaos springing from these dissolving effects of modernization, it is not unnatural that people from traditional religious families look to Islam to provide both a personal feeling of safety and confidence and at the same time social cohesion. However, this call for re-Islamization is a sign that the traditional Islam of the ulema has not kept up with the changes. What the Islamists do is to consciously or unconsciously create a new version of Islam suited to give moral guidance to a changing society while promoting economic progress. Their Islam is marked by a stress on individual responsibility, on the duty of the true believer to engage in a constant struggle against evil within oneself and in society, and to work hard to increase prosperity and welfare. At the same time, it involves new organizational practices, where lay people collect outside the traditional village or kinship groups in associations that study the scriptures to edify the soul and engage in social and political activism for the perceived cause of God.

Islamic Reference in Social and Political Mobilization

On both counts, the relation between the dynamic of re-Islamization, on the one hand, and the dynamics of social modernization and political liberalization, on the other, is far from simply one of opposition. If a reactive rejectionist approach has sometimes been prominent, the complex process of rehabilitation of the vocabulary of Islamic culture seems on the whole rather to extend the field of application of the dynamics of modernization and to accelerate its rhythm than to interrupt it or to reverse its progress. A mobilization centered on questions of identity and morality is rarely confined to particular social groups. And, taken together, existing studies of the social basis of Islamist movements do in fact show a great diversity. Urban workers figure alongside dignitaries of the various bourgeoisies (oil, merchant, military, and tribal), the whole spectrum of middle classes, people from different generations, and those from all spheres of work and levels of education. The factor of gender is often invoked, yet women are well represented along with men in the ranks of those mobilized by Islamism. Within the social diversity involved, the main social stronghold of Islamists, in particular where they remain in opposition, has been and remains among the socially mobile lower-middle and middle classes, especially among students and educated elites. The leading activists have typically been engineers, doctors, teachers, lawyers, and other members of the modern professions along with an important segment of the upcoming business elites. Even if they preserve a "reactive" component inherent in the identity affirmation that is part of the driving force of the movement, the discourse of Islamists and the practices that it legitimates are not only varied but also continuously evolving. Depending on the political context—local or regional—and on the practices of their opponents, they come forth as "literalist" or "liberal," "democrats" or "authoritarian," or "legalist" or "revolutionary."

From the deposed Taliban regime to the Turkish Prime Minister Recep Teyyip Erdogan, the same term *Islamist* has served to designate actors who express themselves within extremely varying doctrinal and political forms. While armed operations are part of the picture for some movements and in some circumstances, parliamentary action has for some time massively been part of the Islamist repertoire of practices. Besides the developments in non-Arab Iran and Turkey, electoral success has led Islamists to leave the pure oppositional position to participate in and even lead governments in Yemen (1993), Jordan (1989), Iraq (2005), and Palestine (2006). For those involved directly in political struggle, it is a distinctly noticeable trend over the past decades and, in spite of the fluctuating levels of repression, toward a more principled stand against the use of violent means to further the cause of Islamic reform and in favor of a democratic political system. Yet again, under the identical label "Islamists," those who are adept at "speaking Muslim" may equally well be members of pietistic movements or new converts who stay mostly outside the political arena, whether of the revolutionary or the legalist kind.

A wide range of discourses, attitudes, and practices may be observed as well in oppositional postures as in the practices of "Islamic" regimes in power. At one extreme of the doctrinal spectrum, the expression of an automated and indiscriminate rejection of any Western democratic institution persists. In that corner, popular sovereignty is considered as strictly incompatible with its divine competitor. At the opposite end, movements that have demonstrated repeatedly their capacity to win a majority of votes (notably in Algeria, Palestine, and in large parts of Egypt) have undertaken a kind of cultural reappropriation of the same popular sovereignty. The rewriting of the modern categories of politics with references from the Islamic repertoire works little by little to heal the symbolic fracture caused by the colonial interruption. Religious affirmation and ethical reaction will certainly not go away. But the religious resource is no longer used to deny the legitimacy of the democratic principle, but on the contrary, it is used more and more to affirm it. The same goes for the defense of human rights (even increasingly, if often ambiguously, women's rights), which have earlier tended to be seen as solely linked to Western culture and thereby disqualified.

Two major movements, the Muslim Brothers and the *salafis*, provide the often diffuse and fluctuating landscape of Islamism with a structuring axis. According to varying national configurations, the fluctuating borderline between the two does not follow exactly the same course in every country or society. The Muslim Brothers and the organizations that have emerged from it, which vary greatly from Tunisia to Kuwait, have for several decades been engaged, where they are given the opportunity, in legal–political action and electoral competition. It is the Muslim Brothers that have contributed the most to popularizing the evolution of classical Islamic thinking and, notably, to the acceptance of an autonomous space for politics and the secularization that is implicit in the recourse to the techniques and the juridical and political categories of the modern parliamentary scene. Where legal recognition continues to be denied to them (as in Tunisia, Libya, and Syria), the activists of the Muslim Brothers engage in educational efforts and devote themselves to work in the associations of civil society and in the antechambers of parliamentary life (professional associations, human rights organizations, and charitable institutions). The principal alternative to the Muslim Brothers is the so-called *salafi* movement, a reference to the "pious ancestors" who are seen to have never deviated from the straight path and by whose example this trend portends to be guided. The *salafi*s have in particular affirmed themselves through a virulent critique of the Muslim Brothers, the main counterimage of their own Islamist identity.

Even if some of the recruits for *jihadi* direct action come from *salafi* ranks, the great majority of *salafis* are pietist and quietist. It is outside the field of armed or even political struggle that they manifest their rupture with society, through a cultural practice by which they wish to be cleansed of all recent or ancient exegetic influences. The *salafis* radically refuse all ideas of a contextualization of the religious norms. They consider the stages in the development of the Muslim Brothers with regard to political liberation and social modernization (and notably the formation of parties or, indeed, organizational structures as such) as so many illegitimate alterations of orthodoxy (*bid'a*). Their doctrine and their practices, less systematically studied than those of the Brothers, who preceded them historically, are neither homogeneous nor static. Conversely, the Muslim Brothers are themselves the object of a modernizing critique coming from more modernist Islamists, who seek to accelerate the very reforming dynamics that the *salafis* want to put a brake on and who denounce the inflexibility and doctrinal rigidity of the Muslim Brothers.

The Islamist regimes, for their part, have contrasting but nevertheless converging performances. Even if the pace is slow, the development of an autonomous political space, the limitation of recourse to repressive violence by the state, and the affirmation of the place of women in the public, professional, and political spheres is hardly moving slower in Sudan, Iran, and especially in the Turkey of the Justice and Development Party than in Algeria, Egypt, or the Tunisia of the "defenders of secularism." In Iran, even if the reference to a double legitimacy, popular and religious, has continued to limit the real importance of the democratic game and the mechanism of parliamentary and presidential elections is far from being definitively sheltered from the authoritarianism of the Supreme Leader and the Guardian Council, the lively debates indicate the formation of a veritable public space. And the possibility offered to the electors of influencing the relation of forces at the top layers of the state is, even if seen against the turbulent aftermath of the "stolen" presidential elections of 2009, rather more credible than that offered to their Algerian, Tunisian, or Egyptian counterparts.

On the scale of the past half-century, it is in fact difficult to demonstrate that the Islamists have in any significant way broken with the norm of institutional and political practices of their secular competitors. Nowhere have they imported violence into a democratic universe, for the simple reason that, in the region under discussion, such a universe has not yet really emerged. Their occasional recourse to political violence may, most often, be correlated with the violence that is exercised against them by forces in the national or international environment. The process of domination seems thus to be at the heart of the principal ideological and strategic trajectories of this category of actors.

François Burgat
IFPO CNRS (Damascus)
Aix en Provence, France

Bjørn Olav Utvik
University of Oslo
Oslo, Norway

See also Church–State Relationships; Civil Society; Democracy: Middle East Perspectives; Fundamentalism; Globalization; Ideology; Islam; Islamist Movements; Mobilization, Political; Nationalism; Nationalist Movements; Religion; Religious Movements

Further Readings

Burgat, F. (1999). *Face to face with political Islam.* London: I. B. Tauris.

Burgat, F. (2008). *Islamism in the shadow of Al-Qaeda.* Austin: University of Texas Press.

Burgat, F., & Dowell, W. (1997). *The Islamic movement in North Africa.* Austin: University of Texas Press.

Eickelman, D. F., & Piscatori, J. (2004). *Muslim politics.* Princeton, NJ: Princeton University Press.

Utvik, B. O. (2003). The modernizing force of Islamism. In F. Burgat & J. Esposito (Eds.), *Modernizing Islam: Religion and the public sphere in the Middle East and Europe* (pp. 43–67). New Brunswick, NJ: Rutgers University Press.

Utvik, B. O. (2006). *The pious road to development: The ideology and practice of Islamist economics in Egypt.* London: Hurst.

Utvik, B. O. (2006). Religious revivalism in 19th century Norway and in 20th century Egypt: A critique of fundamentalism studies. *Islam and Christian-Muslim Relations, 17*(2), 143–158.

Wickham, C. R. (2002). *Mobilizing Islam.* New York: Columbia University Press.

Fuzzy-Set Analysis

Fuzzy-set analysis (FSA) is one branch of qualitative comparative analysis (QCA). The idea of fuzzy sets goes back to earlier insights from computer science and formal logic, which had led to the development of a fuzzy-set approach already in the 1960s. As George Klir, Ute St. Clair, and Bo Yuan (1997) have shown, the difference between fuzzy sets and conventional ("crisp") sets is that set memberships are unequivocal in crisp sets. For example, the element "Sunday" is a member in the set of all days of the week (membership value of 1); whereas the element "January" is not a member of that same set (membership value of 0). By contrast, fuzzy sets have the characteristic that nonperfect set membership can exist. This is especially relevant considering that social science concepts very often have fuzzy boundaries. Democracies are usually not perfect democracies but just democracies to a certain degree. Following this reasoning, most countries have both a partial membership in the set of "democracies" as well as in the set of "nondemocracies." This is also a fundamental difference to crisp sets: In fuzzy sets, one and the same case can have a (partial) membership both in the original set and in the complementary set. A dichotomy is implicitly maintained, but membership degrees render it possible to differentiate for the degree. QCA as a technique is based on the notion of a set. Since crisp-set qualitative comparative analysis (csQCA) can be criticized for requiring dichotomous conditions and outcomes, and since many phenomena in the social sciences are not simply dichotomous, it is understandable why fuzzy sets are attractive for an inclusion in QCA. Below, this entry discusses some of its major features and current developments.

Both the dichotomous version csQCA and fuzzy-set qualitative comparative analysis (fsQCA) are based on an algebra that is different from mainstream linear algebra, as it is applied in conventional statistical methods. csQCA is based on Boolean algebra and fsQCA on fuzzy algebra. However, it is wrong to assume that fuzzy algebra would be an extension of Boolean algebra; rather, Boolean algebra should be seen as a special case of fuzzy algebra. Indeed, all rules for fuzzy algebra are also valid for Boolean algebra but not vice versa. Dichotomies (as used in csQCA and for Boolean algebra) are nothing else than very restricted versions of fuzzy scales, containing just the values of 0 and 1.One of the most difficult (and also most contested) points with regard to FSA is the decision about the attribution of fuzzy values to individual cases. More concretely, if the level of democracy of a given country has to be coded, the question is how the fuzzy values (between 0 and 1) are attributed to single countries, or, in other words, which indicators have to be chosen to attribute a lower fuzzy value to one country and a higher value to another. One possibility is to use already existing quantitative indicators and to standardize them on a range between 0 and 1. However, as Charles Ragin (2000) notes, such a strategy would overlook the central aspect of the coding, namely, that fuzzy values represent the degree of presence of a concept and that they are based on theoretical knowledge. Furthermore, most social science concepts suffer from not having easily definable quantitative equivalents. Therefore, an important part of FSA deals with the translation of qualitative degrees of the presence/absence of a concept into quantitative measures. Indeed, "[i]n the hands of a social scientist [. . .], a fuzzy set can be seen as a fine-grained, continuous measure that has been carefully calibrated using substantive and theoretical knowledge" (p. 7). This also means that in-depth case knowledge is the undeniable prerequisite for a meaningful FSA. Therefore, the coding process usually comes more toward the end of the research process. Recently, Ragin has proposed two more methods to calibrate fuzzy values: a so-called direct and an indirect method that both combine the (theory driven) setting of qualitative anchors with quantitative standardizations in between these anchors (for details, see Ragin, 2008, 71ff).

As in conventional csQCA, the main interest of fsQCA is to find necessary and sufficient conditions (or combinations of conditions) to account for a certain outcome. In fsQCA, the rule is that a condition is necessary if its fuzzy value is always greater than or equal to the fuzzy value of the outcome to be explained. This goes back to the insight that in the dichotomous world a condition is necessary if it exists whenever the outcome is observed. The condition can exist without the outcome being present, but the outcome must not exist without the (necessary) condition being present. This insight from the dichotomous version, namely, that the value of the condition must not be 0 when the value of the outcome is 1, is translated to the fuzzy situation in which values between 0 and 1 are also permitted: A condition is necessary if its fuzzy value is not smaller than the fuzzy value of the outcome. Note that "conditions" in this sense also can mean combinations of conditions that can be created following the rules of fuzzy algebra. Similarly, sufficiency can be defined: A condition is sufficient if its fuzzy value is always smaller than or equal to (but not greater than) the fuzzy value of the outcome to be explained.

Since FSAs very often comprise data for a considerable number of cases and also examine very different conditions, a thorough assessment of sufficiency and necessity cannot be made "by hand." Therefore, a software program has been developed that can be downloaded from the Fuzzy Set/Qualitative Comparative Analysis website. It is based on an algorithm that attributes the individual cases to their (dichotomous) ideal types within which they have various degrees of fuzzy membership values (remember that in fsQCA a case can be a member in more than just one set). These ideal types are then assessed for representing sufficient and necessary conditions and are finally minimized, following the rules of Boolean algebra that are at the heart of each QCA analysis. This analysis is summarized in the so-called Truth Table

Algorithm. In the end, a Boolean equation results (identical to conventional QCA results), from which sufficient and necessary conditions can be extracted. The consistency measure indicates how well the resulting equation corresponds to the empirical data. The overall coverage measure is a measure to indicate which portion of the outcome is explained; this is comparable with the function of an R^2 in quantitative analysis. Raw coverage values show how much each single component of the solution contributes to the explanation of the outcome; unique coverage values indicate which portion of the explanation of the outcome is unique to each single component of the solution term. Consistency and coverage are in a trade-off relationship. It is hard to have optimal values for both of them. Usually, values around 0.7 or (even better) 0.8 are considered acceptable.

As Ragin (1987) notes, another problem of QCA analyses is how to deal with limited diversity—that is, the fact that for some theoretical combinations of conditions no empirical representations (cases) exist. If, for example, a comparative study among Western European countries includes the conditions of a "strong welfare state" and "European Union (EU) membership," we may not find a country that does not have a strong welfare state and is not a member in the EU. It can be argued that limited diversity is an even stronger problem in fsQCA than in csQCA, since not only the two values 0 and 1 are permitted for fuzzy scales but potentially an infinite number of values and therefore also an infinite number of theoretically possible combinations. However, in this case, ideal types are examined. If there are any ideal types to which no single empirical case can be ideally attributed (because the case would be better attributed to another ideal type), then this ideal type is considered a remainder (i.e., not existing in the empirical data), and limited diversity is affirmed. However, it is not limited diversity in the same sense as in conventional csQCA, since each case is a member in all the ideal types up to a certain extent. Too much limited diversity may affect the validity of fsQCA results. Indeed, if a result is based on not enough empirical information, its usefulness can be questionable. There are a number of proposals how to deal with the problem of limited diversity, but of course, no definite answer can be given, since nonexisting data cannot simply be recuperated.

fsQCA can be applied in all social sciences disciplines. An overview is offered on the homepage of the COMPASSS group website. Although it is often considered a small-*N* method, a certain minimum of cases is required. Otherwise, the problem of limited diversity becomes too severe. By contrast, also not too many cases should be examined, since otherwise the intimacy with single cases, which is needed for the calibration of the fuzzy values, is lost. Nevertheless, fsQCA has also been successfully applied to large-*N* data sets.

It is important to underline that fsQCA is not just a technique of data manipulation (the "analytical moment") but that its logic of set-theoretic relationships influences the research process right from the beginning. Therefore, fsQCA should never be applied mechanically but be integrated in a well-prepared research design. Currently, fsQCA scholars work on the improvement of the algorithm and implement ever more features. Additionally, a code of good practice is being developed that should improve advanced applications of the method. As a relatively new technique, fsQCA risks being misinterpreted in many ways.

Claudius Wagemann
Istituto Italiano di Scienze Umane
Florence, Italy

See also Boolean Algebra; Comparative Methods; Configurational Comparative Methods; Qualitative Comparative Analysis

Further Readings

COMPASSS group: http://www.compasss.org

Fuzzy Set/Qualitative Comparative Analysis: http://www.fsqca.com

Klir, G. J., St. Clair, U., & Yuan, B. (1997). *Fuzzy set theory: Foundations and applications.* Upper Saddle River, NJ: Prentice Hall.

Ragin, C. C. (1987). *The comparative method.* Berkeley: University of California Press.

Ragin, C. C. (2000). *Fuzzy-set social science.* Chicago: University of Chicago Press.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond.* Chicago: University of Chicago Press.

Schneider, C. Q., & Wagemann, C. (2007). *Qualitative comparative analysis (QCA) und fuzzy-sets.* Opladen, Germany: Verlag Barbara Budrich.

Index

Entry titles and their page numbers are in **bold.**

International Encyclopedia of POLITICAL SCIENCE

VOLUME 4

Editorial Board

International Encyclopedia of

POLITICAL SCIENCE

VOLUME 4

EDITORS

BERTRAND BADIE
Sciences Po, Paris, France

DIRK BERG-SCHLOSSER
Philipps-Universität Marburg, Germany

LEONARDO MORLINO
LUISS Guido Carli, Italy

SAGE reference

Los Angeles | London | New Delhi
Singapore | Washington DC

IPSA AISP

Developed in partnership with the
International Political Science Association

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
33 Pekin Street #02-01
Far East Square
Singapore 048763

Publisher: Rolf A. Janke
Assistant to the Publisher: Michele Thompson
Developmental Editor: Diana E. Axelsen
Production Editor: Tracy Buyan
Reference Systems Manager: Leticia Gutierrez
Reference Systems Coordinator: Laura Notton
Copy Editor: QuADS Prepress (P) Ltd.
Typesetter: C&M Digitals (P) Ltd.
Proofreaders: Kristin Bergstad, Kate Peterson
Indexer: Mary Mortensen
Cover Designer: Janet Kiesel
Marketing Manager: Kristi Ward

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

International encyclopedia of political science / edited by Bertrand Badie, Dirk Berg-Schlosser, Leonardo Morlino.

8 v. cm.
A SAGE Reference publication.
Includes bibliographical references and index.

ISBN 978-1-4129-5963-6 (cloth)

1. Political science—Encyclopedias. I. Badie, Bertrand. II. Berg-Schlosser, Dirk. III. Morlino, Leonardo

JA61.I58 2011
320.03—dc22   2011009801

11 12 13 14 15 10 9 8 7 6 5 4 3 2 1

Contents

4

List of Entries

Reader's Guide

Comparative Politics, Theory, and Methods

Anarchism
Anarchy
Breakdown of Political Systems
Cabinets
Censorship
Central Banks
Change, Institutional
Charisma
Citizenship
Collaboration
Comparative Methods
Comparative Politics
Competition, Political
Conditionality
Constitutional Engineering
Corporativism
Decentralization
Democracy, Types of
Discursive Institutionalism
Elites
Environmental Issues
Executive
Government
Historical Sociology
Human Rights, Comparative Perspectives
Hybrid Regimes
Institutionalization
Institutions and Institutionalism
Interest Groups
Irredentism
Labor Movement
Leadership
Legitimacy
Military Rule
Monarchy
Neo-Patrimonialism
Neo-Weberian State
Oligarchy
Path Dependence
Personalization of Politics
Pillarization
Political Integration
Political Science, International Institutionalization
Political Systems, Types
Politics of Language
Presidentialism
Prospect Theory
Qualitative Comparative Analysis
Referenda
Reform
Regime (Comparative Politics)
Regionalism
Regionalization
Representation
Republic
Republicanism
Responsibility
Responsiveness
Revolution
Rule of Law
Secession
Semipresidentialism
Separation of Powers
Social Movements
Socialist Systems
Stability
State
State, Virtual
Terrorist Groups
Totalitarian Regimes
Welfare Policies
Welfare State

Case and Area Studies

Area Studies
Authoritarian Regimes

Subnational Governments

Democracy and Democratization

Decision Making in Democracies

Epistemological Foundations

Gender and Race/Ethnicity

International Relations

Political Science, International
Institutionalization
Postmodernism in International Relations
Psychological Explanations of International Politics
Realism in International Relations
Superpower

Peace, War, and Conflict Resolution

Alliances
Arms Race
Bilateralism
Bipolarity and Multipolarity
Civil War
Collective Security
Conflict Resolution
Conflicts
Détente
Diplomacy
Disarmament
Domestic Politics and International Relations
Empire
Foreign Policy Analysis
Genocide
Imperialism
Intervention
Intervention, Humanitarian
Judicial Decision Making
Judicialization of International Relations
Mediation in International Relations
Militias
Multilateralism
National Interest
Natural Resources
Neutrality
Pacifism
Participation, Contentious
Peace
Peacekeeping Positive Peace
Power and International Politics
Preemptive War
Psychological Explanations of International Politics
Sanctions
Secession
Security and Defense Policy
Security Cooperation
Security Dilemma
Sovereignty
Strategic (Security) Studies
Superpower
Territory
Terrorism, International
Transatlantic Relations
Unilateralism
United Nations
Violence
War and Peace
Warlords
Westphalian Ideal State
World Systems Theory

Political Economy

Capitalism
Central Banks
Class, Social
Cost–Benefit Analysis
Economic Policy
Economic Statecraft
Economic Theories of Politics
Foreign Aid and Development
Inequality, Economic
International Monetary Fund (IMF)
International Political Economy
Labor Movement
Market Economy
Market Failure
Monetary Relations
Multilateralism
Multinational Corporations (MNCs)
Nongovernmental Organizations (NGOs)
Policy, Employment
Political Economy
Privatization
Property
Protectionism
Public Budgeting
Public Employment
Public Goods
Redistribution
Social Stratification
Sustainable Development
Tax Policy
Trade Liberalization
Traditional Rule
Tragedy of the Commons
Transaction Costs
Transformation, Economic

Political Parties

Political Philosophy/Theory

Equality and Inequality

Formal and Positive Theory

Theorists

Marx, Karl
Rousseau, Jean-Jacques
Tocqueville, Alexis de
Weber, Max

Political Sociology

Alienation
Anomia
Apathy
Attitude Consistency
Beliefs
Civic Culture
Civic Participation
Corporativism
Credible Commitment
Diaspora
Dissatisfaction, Political
Elections, Primary
Electoral Behavior
Elitism
Empowerment
Hegemony
Historical Memory
Intellectuals
International Public Opinion
International Society
Media, Electronic
Media, Print
Migration
Mobilization, Political
Neo-Corporatism
Networks
Nonstate Actors
Participation
Participation, Contentious
Party Identification
Patriotism
Pillarization
Political Communication
Political Culture
Political Socialization
Political Sociology as a Field of Study
Popular Culture
Power
Schema
Script
Social Capital
Social Cohesion
Social Dominance Orientation
Solidarity
Subject Culture
Support, Political
Tolerance
Trust, Social
Values
Violence

Public Policy and Public Administration

Public Policy

Advocacy
Advocacy Coalition Framework
Agencies
Agenda Setting
Bargaining
Common Goods
Complexity
Compliance
Contingency Theory
Cooperation
Coordination
Crisis Management
Deregulation
Discretion
Discursive Policy Analysis
Environmental Policy
Environmental Security Studies
Europeanization of Policy
Evidence-Based Policy
Immigration Policy
Impacts, Policy
Implementation
Joint-Decision Trap
Judicial Decision Making
Judicial Review
Legalization of Policy
Metagovernance
Monitoring
Neo-Weberian State
New Public Management
Organization Theory
Policy, Constructivist Models
Policy, Discourse Models
Policy, Employment
Policy Advice
Policy Analysis
Policy Community
Policy Cycle
Policy Evaluation

Public Administration

Qualitative Methods

Quantitative Methods

Religion

G

Game Theory

Game theory is the study of rational behavior in situations of interactive decision making—that is, situations in which two or more individuals make decisions that jointly determine an outcome about which the participants have differing preferences or information. As a theory, it aims to articulate the criteria and implications of such decision making, defining principles of idealized strategic choice. These principles can then be used to clarify explanatory and normative concepts of human interaction. An experimental branch examines the extent to which real decision makers in a laboratory setting realize those principles and identifies regular patterns of departure from them. As an empirical tool, game theoretic models of social phenomena yield empirical predictions suitable for testing with observational data. Originally applied to economics, game theory has found application across the social sciences, as well as in biology (where in effect the "selfish gene" and natural selection replace individual strategic decision making) and literary theory (where game theory principles have been used to explicate portrayals of conflict and cooperation, as well as the interaction between author and reader).

Following scattered preliminaries during the period from 1910 to 1927, game theory was essentially created by the mathematician John von Neumann in a paper published in 1928 defining strategic play in finite games of pure conflict with two players. In collaboration with the economist Oskar Morgenstern, von Neumann authored *Games and Economic Behavior* (published in 1944), further developing the theory and beginning its application to problems of economic interaction. Rapid theoretical development followed World War II, and the theory made its first scattered appearances in the political science journals in the 1950s. By the 1980s, game theory had become the primary model-building tool of rational choice political science, with applications in virtually all subfields. International relations saw some of the earliest use of game theory, which came to be especially common in the study of crisis bargaining and the inception of war. The scholars of U.S. politics were also early adopters, applying game theory extensively to legislative processes, electoral competition, and even school budget referendum politics. Many of the American politics applications were generalized to similar questions concerning legislative and electoral politics in comparative perspective, but game theoretic applications to coalition formation, class relations, and ethnic politics were pioneered in comparative politics. In political theory, game theoretic tools were proving useful for gaining new normative understanding in areas such as collective action and constitutionalism.

Game theory itself consists of two main branches: (1) noncooperative and (2) cooperative. Noncooperative game theory focuses on individual decisions, dwelling on problems of strategic moves and incentives, and attempting to explain not only conflict and competition but also the attainment of commitment, coordination, or cooperation in groups. This entry dwells mostly on the noncooperative branch, constructing in detail the two major formulations used there, namely, the *strategic* and *sequential* forms; the major theories of how each form of game is solved; and the correspondences between them. Cooperative game theory focuses on the abilities and actions of coalitions, often assuming that coordination and cooperation can be attained as needed and asking how they will be employed to realize and distribute gains or losses. Cooperative game theory may be taken to subsume not only coalition games but also the foundations of social choice theory and voting theory, although the latter have important overlaps with noncooperative theory as well. The final section summarizes the basic features of the fundamental cooperative game formulation, the *characteristic function* form.

Strategic Games

The strategic game formulation portrays two or more players choosing simultaneously and independently among (possibly different) sets of feasible strategies. Most simply, a *strategy* may represent a simple action: the casting of a vote, for example. More generally, however, each strategy may encompass a complex scheme of conditional actions, reactions, and randomizations, as if chosen entirely in advance and then automatically implemented. The techniques involved in solving strategic games play important roles in analyzing sequential games as well. Political scientists have applied strategic games to a wide variety of political interactions. In some, the choices are, for all intents and purposes, literally simultaneous—secret-ballot voting, for example. More often, the model is used to understand choices with an important simultaneous aspect, such as candidate or party electoral competition and international military confrontations. Strategic games have also proven highly useful in the conceptualization of problems important in political theory and political economy, such as commons problems, free rider problems, and the role of coordination in the formation and interpretation of constitutions.

General Formulation

Three elements specify an elementary strategic game:

1. a set of players, designated for convenience as N or $\{1, 2, \ldots, n\}$, with individual players labeled i, j, and so on;
2. a number of *strategy sets*, one for each player, with i's strategy set denoted S_i; and
3. for each player i, a real-valued *utility function* u_i defined on the set of all outcomes of the game.

For games of *perfect information*, all players are assumed to know the full description of the game: who the players are, what their strategic opportunities are, and what their utility functions are. The players are each to choose, simultaneously and independently, one strategy from their respective strategy sets. One thinks of the resulting *strategy profile* $s = (s_1, s_2, \ldots, s_n) \in S$ as identifying an *outcome*, and the utility function u_i on S represents the preferences of player i. Traditionally, the utility derived by player i when the players choose strategy profile s, $u_i(s)$, is called player i's *payoff* from s. (In many concrete applications, distinct strategy profiles may lead to identical substantive "outcomes," such as "Candidate A wins the election"). Finally, it is often convenient to write s_{-i} for $(s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$, the profile of all players other than i, and to let (s_i, s_{-i}) denote s as well. The utility function, by definition, is meant to capture completely the preferences of a player. This includes any considerations the player may give to attaining some sort of winning margin over other players, to avoiding or seeking risk, or to such considerations as altruism, revenge seeking, and the like. Game theory is not inherently a theory of selfish preferences or material preferences; however, its applications have traction only insofar as correct assumptions about player preferences are employed.

The analysis of some games involves the use of probabilities and expected utility, sources of which are examined below. In the absence of such features, only the players' preference *orderings* over outcomes matters in the analysis of strategic play. If probability is involved, however, then payoffs must be cardinal, in the sense of being von Neumann-Morgenstern utility functions based on preferences over "lotteries" of outcomes. Even so, applying any positive linear transformation of any player's utility function does not affect the analysis of strategic play. Usually, game theory avoids the comparison of payoff values across players for exactly the same reason that such interpersonal comparisons of utilities are eschewed in microeconomic theory: The absolute numbers are taken to have no objective meaning, apart from the ordinal preferences (whether over outcomes or lotteries) from which they are derived. When the number of players and the number of strategies are small, a strategic game can be conveniently described with a *payoff matrix*, which often also gives meaningful labels for the players and strategies in a particular model. For example, Prisoners' Dilemma is often represented as shown in Table 1 where, for example, if Player 1 chooses "Defect" and Player 2 chooses "Cooperate," the payoffs are indicated by the two numbers in the bottom-left cell: 5 for Player 1 and 0 for Player 2. The payoff result in a cell is often written as (5, 0).

Rational Play in Strategic Games

The main concern of game theory is to define what it means for players to choose rationally in such an environment of interactive decision making. A basic concept for defining rational choice is the *best response*: Given that the players other than i choose the strategies s_{-i}, a best response to s_{-i} is a strategy s_i in S_i that maximizes $u_i(s_i, s_{-i})$. If S_i is not finite (as is common in many games of bargaining and of electoral and legislative competition), additional assumptions are needed to ensure existence of a best response. Notice that there may be more than one best response, each giving the same utility. In that case, each best response is called a *weak* but not a *strict best response*. (Note, however, when authors speak of the *set* of player i's weak best responses to s_{-i}, they mean to include the strict best responses as well: that set is a singleton when there is only one best response to s_{-i}, a multiple-element set when there is more than one best response—and the empty set if there is no best response due to an infinite S_i.)

Table 1 A Typical Payoff Matrix: Prisoners' Dilemma

		Player 2 Cooperate	*Player 2* Defect
Player 1	Cooperate	3, 3	0, 5
Player 1	Defect	5, 0	1, 1

The standard formulation of rational strategic play is *Nash equilibrium*. A strategy profile s^* is a Nash equilibrium if and only if for each i, $s_i = s_i^*$ maximizes $u_i(s_i, s_{-i}^*)$ over S_i. In general, Nash equilibria can be identified by deriving the correspondence of best responses of each i to all possible s_{-i},

$$B_i(s_{-i}) = \{s_i \in S_i | s_i \text{ maximizes } u_i(s_i, s_{-i})\}$$

and then finding the intersection of those correspondences, that is,

$$\{(s_1, s_2, \ldots, s_n) | \text{for each } i, s_i \in B_i(s_{-i})\}.$$

This definition satisfies two criteria for strategic rationality. First, each player chooses a best response to the choices he or she anticipates the other players will make, while anticipating that the other players too are strategically rational. This criterion alone defines a weaker solution concept, a *rationalizable* strategy profile. As we will see, however, this may or may not result in outcomes in which each player has chosen a best response to the actual choices of all other players.

Second, Nash equilibrium satisfies a criterion of what one might call rational *experienced* play. This criterion is embodied in the assumption that each player's anticipation of the others' choices actually matches those choices. Suppose groups of n individuals in a population of much larger size are randomly combined from time to time to play a single iteration of the original game and that all have information about how the game is being played. Each individual has a contingent strategy plan—that is, a plan of which strategy to choose if placed in the role of Player 1, Player 2, . . . , or Player n. Suppose further that individuals revise their plans when they observe that an alternative strategy could have yielded a higher payoff to profiles recently played than did the player's previous intended choice and that those revisions tend to adopt higher payoff strategies against the observed actions of others. In this setting, any population strategy profile that does not induce a Nash equilibrium for every possible assignment of n individuals to the player roles in the game will be subject to such change by some player. That is, only Nash equilibria can be stable. (The literature on *learning in games* formalizes this intuition.)

Coalitionally Rational Play

The ideas of best response and hence of Nash equilibrium imagine, in effect, that every player considers only individual deviations from a profile. For equilibrium profiles in some games, however, it may be the case that two or more players could all have achieved higher payoffs had they all chosen differently. The analysis of possible coalitional action in a strategic game can sometimes be accomplished through a version of Nash equilibrium that inquires not only about individual departures from a strategy profile but also about simultaneous departures by several players. For any strategy profile s and any subset or *coalition* of players $C \subseteq N$, write s_C for the list of strategies chosen by players in C, s_{-C} for the list of strategies chosen by players not in C, and let (s_C, s_{-C}) also denote s. A strategy profile $s^* = (s^*_C, s^*_{-C})$ is a *strong Nash equilibrium* if for every coalition C and for every feasible alternative list of strategies s_C (where for each $i \in C$, $s_i \in S_i$), there is some member j of C for whom

$$u_j(s_C, s^*_{-C}) \leq u_j(s^*_C, s^*_{-C}).$$

That is, for no coalition is it possible to find an alternative profile for coalition members such that all would have been strictly better off had all chosen those alternative strategies, given the strategies of nonmembers. Note that a strong Nash equilibrium is a Nash equilibrium, since the coalitions in the definition must include single-element sets of players, $C = \{i\}$. However, not every Nash equilibrium is a strong Nash equilibrium. Moreover, no strongly inefficient strategy profile can be a strong Nash equilibrium, since the coalitions in the definition must include the universal coalition, $C = N$.

For example, in the Prisoners' Dilemma, (Defect, Defect) is a Nash equilibrium but not a strong Nash equilibrium, since all members of the universal coalition $C = \{\text{Player 1, Player 2}\}$ could have achieved a higher payoff had they instead chosen (Cooperate, Cooperate). The latter profile, however, is not a strong Nash equilibrium either, since the singleton coalition $C = \{\text{Player 1}\}$ could strictly improve its members' payoffs by instead choosing Defect. The Prisoners' Dilemma has no strong Nash equilibrium.

Mixed Strategies and Existence of Nash Equilibrium

A game may not have any Nash equilibria in such deterministic (or "pure") strategies. A classic example is the game usually called "Matching Pennies," shown in Table 2. More generally, this happens in games in which one player wants, in some sense, to match the other but not vice versa: Rock-Paper-Scissors and Soccer Penalty Kicks are also common examples. Such examples are not confined to games of pure conflict, however. In such situations, it behooves a player to be as unpredictable as possible to his or her opponent. Game theorists denote this through the device of a *mixed strategy*, in which player i randomizes over S_i. In an analysis allowing mixed strategies, a strategy is denoted by a probability distribution over S_i. (The set of all mixed strategies includes all the deterministic strategies as well, which are denoted by degenerate probability distributions placing all weight on one element of S_i.) And in that setting, players are assumed to choose strategies so as to maximize *expected payoff*.

With the addition of mixed strategies, this seminal existence result known as *Nash's theorem* holds: Every finite game has at least one Nash equilibrium in mixed strategies.

Multiple Nash Equilibria

Nash equilibrium eliminates some rationalizable strategy profiles, but a game may still have many Nash equilibria. The canonical example is the game of pure coordination, shown in Table 3, which has two pure-strategy (and one mixed

Table 2 A Game With No Pure-Strategy Equilibria: Matching Pennies

		Player 2	
		Heads	Tails
Player 1	Heads	−1 1	1 −1
	Tails	1 −1	−1 1

Table 3 A Game With Two Pure-Strategy Equilibria: Pure Coordination

		Player 2	
		A	B
Player 1	A	1, 1	0, 0
	B	0, 0	1, 1

strategy) equilibria. Without additional assumptions—concerning, for example, iterated play and dynamics—game theory proper has nothing further to say about which equilibrium would be chosen. In applications, the historical or cultural context in which a game is played might afford the players additional information about what each other will anticipate, resulting in a *focal point*, a choice among equilibria that players can arrive at more often than random chance or symmetric mixed strategies would predict. Games with multiple equilibria demonstrate the difference between rationalizable and Nash strategy profiles: If Player 1 chooses A based on the belief that 2 will choose A, and Player 2 chooses B based on the belief 1 will choose B, this is a rationalizable profile of strategies and beliefs. But the result disconfirms the players' beliefs and disappoints their expectations.

Evolutionary Game Theory and Learning in Games

Evolutionary game theory offers an additional set of assumptions and tools to say more about the conditions under which various Nash equilibria will be selected. These tools are used in biology to study actual evolution through natural selection (usually with "selfish genes" filling the role of rational players). Individuals whose phenotypes—strategies—yield them higher reproductive fitness in one generation have more progeny in the next. A distribution of genotypes stable in the resulting dynamic system corresponds roughly to equilibrium in the game.

In the social sciences, the tools of evolutionary game theory are more commonly applied to models of "learning in games," in which individuals play games many times against different opponents, occasionally switching strategies to adopt a new one observed to have been more successful. The abandoning of one strategy in favor of another takes the place of the biologist's process of natural selection. This learning process can be quite myopic and error ridden, but generally the result is that a population of individuals may reach a *stable point*, a whole-population profile of strategies in which little further systematic improvement is possible. Stable points correspond under general conditions to a subset of Nash equilibria. Models of learning in games are significant to game theory overall because they demonstrate that the strong assumptions used to construct basic games—perfect and common information, perfect ability to select optimal strategies, and an a priori coincidence in the selection among multiple equilibria—are not remotely necessary to produce choices that correspond to classic Nash equilibrium. What is required is a sufficient opportunity for players to gain experience in playing.

Imperfect Information in Strategic Games

The assumption of perfect information is obviously an extreme idealization of strategic interaction. Even in classical game theory (as opposed to evolutionary or population learning models), a major preoccupation is the generalization of this ideal case. This is accomplished by specifying probability distributions that describe the state of a player's knowledge, including knowledge about the state of other players' knowledge. In general, imperfect information about the set of players, the strategy sets, or the utility functions in a strategic game can all be expressed in terms of uncertainty about the utility functions alone: An unavailable strategy can be represented by an extremely low utility for any profile in which that strategy is used, and an absent player can be assigned a singleton strategy set. Often, the true payoff function of a player is called the player's *type*.

Once each player's imperfect information is expressed in the form of a probability distribution over payoffs, the criterion of best response can be expressed in terms of expected-payoff maximization. Nash equilibrium generalizes directly to this form of best response. An important complication is that, in calculating best responses, a player's subjective probabilities about other player types must be updated to account for the circumstances under which the first player's action will affect his or her payoff. For example, in models of instrumental voting, only the cases in which a player's vote will be *pivotal* have any importance for maximizing payoff; but the fact of being pivotal carries implications about what the other players' types might be in those circumstances.

Sequential Games

Strategic games abstract away players' opportunities to react to other players or to anticipate their future decisions, burying such activity implicitly within each player's strategy. Sequential games, by contrast, offer an alternative model designed to emphasize this aspect of interactive decision making. Political processes are usually sequential decision processes, and political scientists have exploited this in numerous ways. Crisis bargaining and the inception of war were early and important objects of study through sequential games; later models of ethnic conflict and cooperation involve an approach similar in spirit. In the study of policy formation through voting, agenda-setting models such as committee gatekeeping have been widely employed, and the theory of legislative bargaining has achieved the status of a general-purpose tool for legislative politics, political parties, and group choice generally. Sequential game models are prominent also in the study of delegation of powers to an agency or committee and of decision making within a system of separated powers. Games of sequentially repeated interaction are often applied to the formation and maintenance of community or of collective action, in settings ranging from the theoretical study of polities to the stability of international organizations. With the incorporation of incomplete information, sequential games have also found important applications to all manner of political communication phenomena, such as policy advice from biased experts, lobbying, and informative political debate.

General Formulation

The following elements are necessary to specify a basic sequential game; for special purposes, additional elements may be needed, and these will be described in the subsequent sections. Just as with strategic games, every sequential game has a set N of players and, for each player i, a utility function u_i over outcomes. The difference lies in the description of player decisions and, thus, of outcomes. A sequential game also consists of a description of turns on which a given player has the chance to act and the actions among which that player may choose. A game need not simply describe the turns as alternating regularly among players; in general, the coming of a player's turn, as well as the actions available to that player, may depend on the prior actions of all players. The description of the game must also fully specify the sequences of player actions that determine the various outcomes of the game. As with strategic games using this description, any outcome may be simply specified by the sequence of actions that led to it, although in applications the outcomes usually correspond to substantive results.

Traditionally, and in most small games, the whole description of turns, actions, and outcomes is accomplished through the specification of a *game tree*. The "tree" terminology refers to a *rooted tree* in the mathematical theory of graphs. In that theory, a *graph* consists of a set of *nodes* and the *edges* that connect some nodes directly to other nodes; a *tree* is any graph that is both *connected* (any node can be reached from any other node via a sequence of edges that does not retrace itself, called a *path*) and *acyclic* (has no cycles, i.e., no path from a node back to itself); and a *rooted* tree has one node designated as the *initial* node (or *root*). The properties of connectedness and acyclicity then imply that for any two paths leading away from the initial node, once those two paths diverge onto different edges they never reconnect—giving the graph the shape, schematically, of the branches of a tree.

The initial node represents the beginning of the game. The path leading from the initial node to any particular node represents a partial or complete *history* of play of the game (including the "null history" at the initial node); at each node along any path from the initial node, either it is one player's turn to move or else the path outward from the root ends. These nodes along any path are called *nonterminal* and *terminal* histories of the game, respectively. The edge connecting the node of a nonterminal history, h, to the next node outward from the root represents an *action* chosen by the player $P(h)$ whose turn came at h. Each terminal history represents an *outcome* of the game, the result of a particular *path of play*.

This description of a sequential game as a game tree is just a way of forcing the specification of turns, action choices, potential histories, and outcomes to satisfy appropriate regularity conditions, making the description unambiguous and internally consistent. With the addition, described below, of chance moves, simultaneous moves, and information sets, the structure is remarkably versatile and yet amenable to powerful proofs concerning the general properties of types of games and their solutions. Figure 1 shows an example of a simple game tree, a so-called entry or incumbency game. In the figure, the payoffs of the two players are shown by a label on each terminal history x, with $u_1(x)$ written above $u_2(x)$. This is sometimes written instead as $(u_1(x), u_2(x))$.

Note, however, that the general form of a sequential game may involve trees that cannot be literally portrayed in a tree diagram such as Figure 1. For many applications, analytic tractability is improved by allowing a player i at a given history h to have a set of available actions $A_i(h)$ that is infinite, as in games of bargaining over the division of a stock or in a spatial model of electoral competition. Also, it is often important that there be no a priori limit on the possible length of a path of play, so that some terminal histories are infinitely long, as in alternating-offer bargaining processes or in repeated games. A well-specified model must somehow limit the payoff to such an infinite "terminal" history (e.g., by a low reversion payoff in bargaining or by discounting in a repeated game). An infinitely long path of play is nevertheless, of course, unrealistic as an outcome; but the ever-present possibility of continuing the game is strategically critical. A game having at least one terminal history of infinite length is said to have "infinite horizon." A game having no such history, and in which the set of available actions at every turn is finite, is a "finite game."

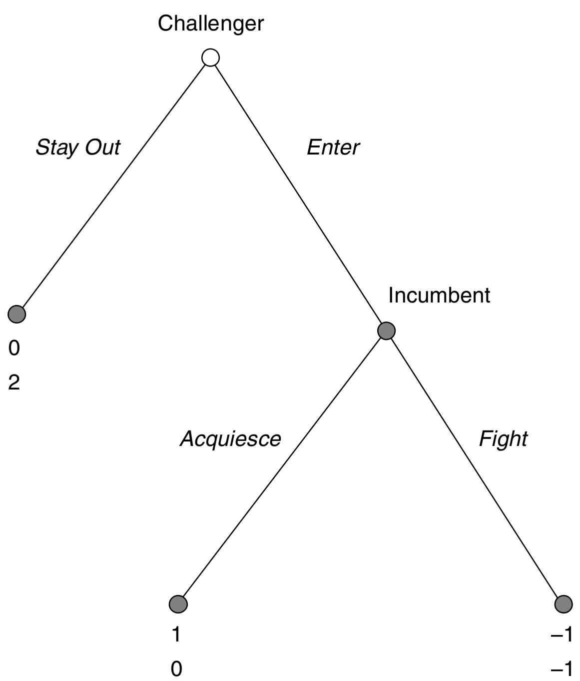


Figure 1 A Typical Game Tree: The Entry Game

Strategies and Solutions in Basic Sequential Games

A strategy for player i in a sequential game consists of a *function* s_i specifying, for each h having $P(h) = i$, an action $s_i(h)$ from i's set of available actions at that history. A strategy profile is a list of such functions, one for each player. As with strategic games, a strategy profile is written equivalently as s, $(s_1, s_2, \ldots, s_n)$, or (s_i, s_{-i}).

Backward Induction

For any finite basic sequential game, a simple and appealing solution concept is available: the method of *backward induction*. Consider an "ultimate" nonterminal history, that is, a history h such that every action taken by player $P(h)$ will end the game. Since there is a finite number of such actions, at least one of them yields a maximum payoff. Here at the end of the game, the only relevant implication of this player's choice is to determine his or her own payoff; clearly the rational choice is one yielding that maximum. Backward induction begins by identifying one such payoff-maximizing action for every ultimate nonterminal history. Now consider a "penultimate" history: a history h' for which (h', a') is an "ultimate" history no matter which action a' is chosen by player $P(h')$. For a given a', let a be the action that will then be taken at the resulting ultimate history (h', a'). In that case, using our analysis of ultimate histories, player

$i = P(h')$ will receive a predictable payoff from a', namely, u_i (h', a', a). Proceeding in this fashion, backing up the tree to the initial node specifies an action choice for the choosing player at every nonterminal history, defining a strategy for every player. This strategy profile defines the sequence of actions that will transpire when the game is played—a terminal history or complete path of play through the game tree—as well as the actions that would be chosen following any arbitrary departure from that path. This strategy has the property that all actions chosen, both on and off the path, are payoff maximizing, given all the subsequent choices by all players. Call this strategy profile the *solution by backward induction*.

Nash Equilibrium: The Strategic Form of a Sequential Game

Alternatively, a sequential game can be converted to an essentially equivalent *strategic form* and solved by specifying a Nash equilibrium. For a finite basic sequential game, consider the strategic game defined by (a) the same set of players *N*; (b) for each player *i*, the strategy set consisting of all full sequential-game strategies for player *i*; and (c) for every profile of such strategies, a payoff equal to the sequential game payoff corresponding to the terminal node that would result from applying that strategy profile to the sequential game. A Nash equilibrium of this strategic form is defined as a Nash equilibrium of the original sequential game; the strategic form is also known as the *normal* form or occasionally the *matrix* form. The sequential form is also known as the *extensive* form or occasionally the tree form of the game.

Notice that every solution by backward induction is a Nash equilibrium of the game. In a backward-induction solution, no player has any incentive to change *both* an ultimate and any penultimate moves because all the ultimate moves are optimal regardless of how those ultimate histories were arrived at, and the penultimate moves were optimal given those. Backing up through the game tree, similar reasoning establishes that no player would want to change any subset of his or her moves in the backward-induction solution; hence the profile is a Nash equilibrium.

Not every Nash equilibrium is a backward-induction solution, however. Consider the Challenge Game shown in Figure 1. Backward induction yields a profile in which Incumbent Acquiesces, and Challenger Enters, which is also a Nash equilibrium. Consider another strategy profile in which Challenger chooses Stay Out and Incumbent chooses Fight. Clearly, Challenger would not prefer to Enter if Incumbent will Fight; moreover, given that Challenger will Stay Out, Incumbent has no incentive to Acquiesce. This too is a Nash equilibrium, but it violates backward induction. Such a flawed Nash equilibrium is sometimes called "imperfect."

Subgame-Perfect Equilibrium

A subgame of a sequential game *G* is the part of the game that follows any nonterminal node *h*, and thus, it is a game in its own right. Its set of histories corresponds to the set of all histories of *G* that begin with *h*; its payoffs are the corresponding payoffs from *G*; and its player set is the same, although some may have no moves to make in the subgame. (For games of imperfect information, defined below, this definition is modified to exclude putative subgames whose set of nonterminal histories would contain an incomplete portion of any information set.) When a game is solved by backward induction, a backward-induction solution on every subgame is induced as well. Thus, backward induction induces a Nash equilibrium on every subgame.

This property can be applied as well to a game with infinite horizon, which cannot be solved via backward induction. For any sequential game, a Nash equilibrium that induces a Nash equilibrium on every subgame avoids the "imperfection" described above; hence, it is called a *subgame-perfect equilibrium*. Since not every Nash equilibrium is subgame perfect, subgame perfection is also referred to as a *refinement* of Nash equilibrium. (The qualification "subgame" perfect distinguishes this refinement from other refinements treating alternative notions of imperfection.) Where backward induction is not possible or feasible, other methods may be employed to derive subgame-perfect equilibria. Such methods are often based in one way or another on Bellman's principle of dynamic programming: An optimal strategy is identified by always taking the optimal current action, assuming that the subsequent actions are optimal. Other Nash equilibrium refinements in a similar spirit use criteria based on rational behavior when choices are known to be subject to small errors or *trembles*, both in sequential and in strategic games. Examples include *trembling-hand perfection* and *proper equilibrium*.

Extending the Basic Formulation

The basic formulation of a sequential game, whether finite or infinite, can be extended in several ways. Backward-induction and subgame-perfect equilibrium extend as well, making the sequential game a powerful all-purpose modeling tool. This section examines how to generalize the formulation of sequential games to include simultaneous moves, chance occurrences, randomized actions, and imperfect information.

Simultaneous Moves

Suppose that, following some histories, two or more players are to choose actions simultaneously, after which the game may continue. This can be represented by letting the player function P designate a set of players, $P(h) \subseteq N$, rather than a single player. For each $i \in P(h)$, we still let $A_i(h)$ represent the set of actions among which i is to choose; this may, in general, be a different set of actions for each player in $P(h)$. The description of a history accommodates such simultaneous moves by allowing the sequence of actions defining h to include, where appropriate, a profile of actions by a set $P(h)$ of players, in place of the move by a single player. For example, if action a^1 by Player 1 is followed by a simultaneous choice by Players 2 and 3, who choose a^2 and a^3, respectively, then resulting history may be written unambiguously as $(a^1, (a^2, a^3))$. Player i's strategy in a sequential game can still be represented in the same form as before, letting $s_i(h)$ represent the action i would take at history h regardless of whether other players are simultaneously making choices at h. A basic sequential game is one in which $P(h)$ is always a singleton $\{i\}$. A strategic game now becomes a special case of a sequential game, in which the null history $\varnothing$ is the only nonterminal history, with $P(\varnothing) = N$.

Finally, it is natural to extend backward induction to simultaneous-choice situations as follows: when, in the process of backward induction, a history h is reached at which $P(h)$ has more than one player, then construct a special strategic game in which the set of players is $P(h)$; player i's strategy set is $A_i(h)$; and player i's payoff to strategy profile $s(h)$ in the special strategic game is identical to i's payoff in the overall sequential game from the terminal history defined by h, followed by $s(h)$, followed by continued play according to the backward-induction steps already completed. The backward induction is continued, then, by designating a Nash equilibrium $s^*(h)$ in the special strategic game at h, and then continuing up the tree under the assumption that if h is reached, the players in $P(h)$ will choose according to $s^*(h)$. As in backward induction in a basic sequential game, if a simultaneous-move history has more than one Nash equilibrium, the implications of each must be traced out for the remaining steps of backward induction, yielding several different backward-induction solutions. Extending the method of backward induction in this manner preserves the definition of strategic form, the definition of Nash equilibrium for sequential games, and the correspondence between backward-induction and subgame-perfect equilibrium all in the same form as defined for basic sequential games.

Example: Repeated Games

An important application of sequential games with simultaneous moves is the case of *repeated games*. Let G be a strategic game (players N, strategy sets S_i, and utility functions u_i). In a repeated game, the same set of players plays G, known as the *stage game*; all learn the outcome; and all play G again, and so on. The players know that G will be repeated a specific number of times or indefinitely. The payoff to this entire sequence is the sum of the payoffs to the successive instances of G, perhaps discounted (i.e., multiplied by some "discount factor" between 0 and 1 for each repetition) due to time preferences or to the probabilistic termination of the sequence.

Consider for instance the Prisoners' Dilemma shown in Table 1 as the stage game, to be repeated infinitely with discount factor w. This forms an infinite-horizon sequential game, in which at every nonterminal (i.e., every finite) history h, $P(h) = \{\text{Player 1, Player 2}\}$ and $A_i(h) = \{\text{Cooperate},$

Defect}. A terminal history is an infinite sequence of the form $((a_{10}, a_{20}), (a_{11}, a_{21}), (a_{12}, a_{22}), \ldots)$, where each $a_{it} \in$ {Cooperate, Defect}. The payoff to player i from such a terminal history is

$$\sum_{t=0}^{\infty} w^t u_i(a_{1t}, a_{2t}),$$

where t represents time.

Finally, a strategy takes the form of a rule telling what action would be chosen at every possible nonterminal history. Actual strategies of interest are usually expressed in the form of a relatively simple rule. For example, the strategy of Tit for Tat in the repeated Prisoners' Dilemma can be described as follows: Cooperate at $t = 0$; and for all $t > 0$, imitate the other player's action at $t - 1$.

A repeated-game strategy profile can be proven to be a subgame-perfect equilibrium by showing that at no possible history (on or off its path of play) could payoff be increased by changing the action taken at that history and nowhere else. This method is derived from Bellman's principle. Notice that the profile in which both players use Tit for Tat is a Nash equilibrium for sufficiently large w but is not a subgame-perfect equilibrium for any value of w. This is because if one player ever commits an unexplained Defection, then, under the assumption that both players would adhere to Tit for Tat in the future, it would be better for the other player to forgo punishing the defection (incurring a punishment in turn). If the strategy for both players is modified, however, to include an "apology"—namely, Cooperate twice regardless of the other player's action—for any unprovoked Defection, it becomes a subgame-perfect equilibrium (again, for sufficiently large w). Thus, although Defect is a dominant strategy in the unrepeated game, full cooperation can be maintained in equilibrium in the indefinitely repeated game.

Moves by Chance

The interactions we subject to game theoretic analysis sometimes depend in part on intervening events best understood by the players as being random. Such phenomena can be included by allowing the player function to specify that at some histories h, instead of a player or players choosing an action, a random event occurs before subsequent player actions or payoff determinations take place. This is denoted by writing $P(h) =$ Chance or $P(h) =$ {Chance} and such an h is casually referred to as a "move by Chance" or by "Nature." In describing the game, any such instance must be accompanied by a set of possible outcomes of the random event, which we might as well write $A_{\text{Chance}}(h)$, and a probability distribution p on $A_{\text{Chance}}(h)$. The outcome of a Chance move is part of the sequence of actions defining any history.

If Chance moves are encountered in the path of play under a profile s, then s determines not a particular terminal history but rather a probability distribution over terminal histories. For a game involving moves by Chance, then, the backward-induction process must be modified by maximizing *expected* payoffs, computed in the appropriate way using the probability distributions for the Chance moves to determine the probabilities of various payoffs given a strategy profile.

Behaviorally Mixed Strategies

Mixed strategies are usually treated in a different way for sequential games than for strategic games. Game theorists use the term *behavioral* (or *behaviorally*) *mixed strategy* to refer to a strategy in a sequential game in which the player randomizes over the actions. (Where the context is clear, the modifier "behavioral" is usually dropped.) Unless otherwise specified, these randomizations are always taken to be independent across players; they are represented as independent across histories too, although of course different actions at an earlier history lead to different subsequent histories, at which the player may use different randomizing probabilities.

Obviously, a strategy profile involving randomization will determine a probability distribution over outcomes, and expected utility is the criterion for backward induction. Analogous to deriving mixed-strategy equilibria in strategic games, the procedure in a sequential game for finding subgame-perfect equilibria involving behavioral mixed strategies depends on the fact that a player who maximizes expected utility can rationally randomize if and only if each of the actions accorded positive probability would, if used alone, also maximize expected utility. In the backward-induction process, this means that a single player $P(h)$ could use either no randomized action or a continuum of randomized actions at h. However, many of these possibilities will often be ruled out of the eventual subgame-perfect equilibrium, since randomizations by more than one player or at more than one history will require multiple indifference conditions that (again as in strategic games) constrain the feasible values of the probabilities used.

Imperfect Information

Perhaps the single most important extension of the basic formulation is to represent games in which players do not have full information about the previous actions of other players. This is, after all, an important feature of many parlor games as well as a key feature of real-world interactions. The state of player i's knowledge about previous moves in a sequential game is represented by distinguishing, as an *information set* of player i, any subset of histories at which player i is required to move but among which i is unable to distinguish at the time they are reached. Thus, for all histories within the same information set, the player is required to choose the same action. The whole collection of such information sets for a player is that player's *information structure*. To define equilibrium in a game with imperfect information requires that we incorporate a notion of player i's probabilistic beliefs about previous actions. These notions are formalized below to facilitate defining the relevant equilibrium concepts.

For each player i, let $H(i)$ be the set of histories for which i is in the player set. Let I_i be a partition of $H(i)$—that is, a set of nonempty subsets of $H(i)$ such that each history in $H(i)$ is an element of exactly one of those subsets. Denote by $I_i(h)$ the partition element containing h, called i's *information set* at h. In what follows, the symbol Z denotes a typical information set without reference to a particular history in it; but for any h in Z, $I_i(h) = Z$. I_i is a valid *information structure for i* if and only if, for each information set Z in I_i, player i has exactly the same action choices at every h in Z. (In any application, the names or labels of the actions for each history in an information set should be identical as well. From a theoretical standpoint, however, the labels on the actions are irrelevant; hence, the only real restriction is that the action sets be of the same size.) This restriction is needed since i must be aware of the available choices at the present history; it must be impossible for the player to infer from this whether that history is h or h'. Thus, we may as well write $A_i(I_i(h))$ for the actions available at every h in information set $I_i(h)$. Finally, for any history h at which player $i \in P(h)$ does know the entire sequence of previous actions, $I_i(h) = \{h\}$, just that one history.

This approach requires a slight redefinition of a strategy. Since player i cannot distinguish between h and h' in an information set Z, a strategy must specify the same action at both histories. Thus, in general, a strategy for player i can be fully described as a mapping not from each history into its action set but simply from each information set $Z \in I_i$ into its action set; thus, instead of $s_i(h)$, we keep track of the accounting by writing $s_i(Z)$, which automatically reminds us that $h, h' \in Z$ implies $s_i(h) = s_i(h')$. As noted above, the presence of information sets restricts the applicability of the notion of subgame-perfect equilibrium. In a proper subgame, every h within the subgame must have the property that the subgame contains *all* the histories in $I_i(h)$. That is, no information set can be "shared" between two subgames. Thus, in a game of imperfect information, not every nonterminal h gives rise to a subgame. With this restriction on what qualifies as a subgame, the definition of subgame-perfect equilibrium remains the same as before.

To go further in assessing the choice of a rational strategy requires that we devise some method of handling a player's uncertainty at any nontrivial information set Z. As usual, we represent the state of the player's knowledge about what is the actual history in Z with a probability distribution over Z. Now, the payoff from choosing action a from $A_i(Z)$ is the expectation over the histories in Z of the payoffs from choosing a. Each of these payoffs, in turn, may be an expectation based on subsequent moves by nature or randomized actions to be chosen by i or by other players subsequent to that. We thus wish to equip each player i with a probability distribution over each nontrivial Z in I_i. Let $\mu_Z(h)$ represent i's subjective probability that the current history is h given that the current information set is $Z \in I_i$. The distribution μ_Z is known as player i's *beliefs* at Z, and μ as defined over all Z for all players is a *system of beliefs*. Using this notion of beliefs, we can generalize subgame-perfect equilibrium to the imperfect-information setting. There are several such extended definitions, each having the following form: An equilibrium consists of an *assessment*, that is, a profile of strategies s and a system of beliefs μ, obeying two conditions:

1. *Sequential rationality*: For each i and each $Z \in I_i$, $s_i(Z)$ maximizes the expected value of u_i given s_{-i} and μ.
2. *Consistency*: Each μ_Z is consistent with s and with the probabilities placed on Chance moves, given the laws of probability and other considerations.

The sequential rationality condition is just a restatement of the criterion we have been using all along, that each action maximizes expected payoff. The consistency condition varies according to the "other considerations" imposed as conditions, and this gives rise to different solution concepts.

The simplest consistency condition—and the least restrictive while still being consistent with the laws of probability—is that beliefs must comport with Bayes's rule wherever it is directly applicable. That is, for every information set Z that can actually be reached with positive probability given s, for every history h' in Z, we must define $\mu_Z(h')$ according to Bayes's rule, namely, as the conditional probability of reaching h' given that Z is reached, given the players' strategies and the game's Chance-move probabilities. This amounts to

$$\mu_Z(h') = \Pr(h')/\Pr(Z) = \Pr(h')/\Sigma_{h \in Z} \Pr(h),$$

for all h' in Z. Each marginal probability $\Pr(h)$ is calculated by multiplying all the mixed-strategy action probabilities and Chance-move probabilities along the sequence of actions that make up h. An assessment (s^*, μ^*) satisfying this form of consistency, as well as sequential rationality, is a *weak sequential equilibrium* (WSE).

Game theoretic models frequently invoke a related equilibrium concept called *perfect Bayesian equilibrium* (PBE). A sequentially rational assessment (s^*, μ^*) is a PBE if and only if it induces a WSE on every subgame. The only difference between PBE and WSE occurs when there is an information set Z that is not reached by s but for which other off-path actions or chance moves intervene between the point where s diverges from Z and Z itself. In such cases, WSE consistency imposes no restriction on beliefs in Z, but PBE consistency requires that the probabilities in Z comport with the probabilities of reaching Z conditional on the initial departure from the path of s. Other, stronger consistency conditions produce further *belief-based refinements* of WSE, such as sequential equilibrium and divine equilibrium. In general, such refinements are used to rule out equilibria thought to be based on "unreasonable" beliefs.

Example: Signaling Games

The important example of "semipooling" equilibrium in certain *signaling games* provides important applications of chance moves, behavioral mixed strategies, and imperfect information. In a signaling game, a Sender learns the value of an initial Chance move and then sends one of a set of feasible signals to a Receiver, who must choose an action. The players' payoffs depend in different ways on the combination of the true chance move and the Receiver's move. If the Sender's payoff does not depend directly on the Sender's signal, the game is one of "cheap talk"; otherwise it may be a game of "costly signaling." Depending on the details of the conflict between the players' interests, in equilibrium the Sender's signal may be (a) ignored by the Receiver, called "pooling" since the Sender's behavior is the same regardless of the Chance move; (b) perfectly informative about the Chance move, called "separating" since the Sender's signal perfectly distinguishes the possible Chance moves; or (c) partially informative, called "semipooling" since the Sender uses a mixed strategy, reporting truthfully with some probability less than 1.

Figure 2 shows the game tree for a signaling game that minimally embodies all these features. The game begins (in the middle of the diagram) with Chance determining whether Policy A or Policy B would be best for Receiver; A is best with probability $p < .5$. Sender learns the truth and signals either a or b. Receiver knows p but does not know the Chance move; thus, Receiver has two nontrivial information sets, one reached when Sender signals a and the other when Sender signals b. These are indicated in the game tree by the dotted lines connecting histories that Receiver cannot distinguish. After hearing the report from Sender, Receiver must choose either α (implement Policy A) or β (implement Policy B). At each terminal history, each pair of payoffs in the diagram shows Sender's payoff above Receiver's payoff. Thus, Receiver's preferences simply favor Policy A when A is best and B when B is best, with equal intensity. We assume that Sender's payoff parameters obey $x > w > y > 0$, so Sender has a preference for A regardless of which policy is best for Receiver. Since $w > y$, Sender's message has a direct impact on Sender's utility: if B is best and B is implemented, Sender's utility will be reduced (by $w - y$) for having falsely reported that A is best. This is, therefore, a game of costly signaling rather than cheap talk.

This game has a semipooling equilibrium defined by the following assessment, which incorporates a behavioral mixed strategy:

- Sender signals a when A is best; when B is best, Sender signals a with probability $r = (1 - 2p)/(1 - p)$ and b with probability $1 - r$.
- Following a signal of b (the left-hand information set), Receiver believes that B is best with probability 1. Following a signal of a (the right-hand information set), Receiver believes there is a probability $m = p/[p + (1 - p)r]$ that A is best and a probability $1 - m$ that B is best.
- Following a signal of b, Receiver chooses β with certainty. Following a signal of a, Receiver chooses α with probability $s = (w - y)/(x - y)$ and chooses β with probability $1 - s$.

To see how the belief m is derived from Bayes's rule, notice that the top history in the right-hand information set is reached with probability p (Chance probability that A is best, times Sender's probability 1 of truthfully reporting that) and the bottom history is reached with probability $(1 - p)r$ (the Chance probability B is best times Sender's mixed-strategy probability of falsely reporting A is best). Notice that this formulation of m is equivalent to the usual form of Bayes' rule: m is the conditional probability that A has occurred, given that Sender has reported a, so Bayes' rule says $\Pr(A|a) = \Pr(A)\Pr(a|A) / [\Pr(A)\Pr(a|A) + \Pr(B)\Pr(a|B)]$, which is just $p / [p + (1 - p)r]$, as given above.

As with any signaling game, there are also pooling equilibria in which Sender uses the same signaling probabilities regardless of Chance's move; and Receiver chooses β in all cases, believing B is best with probability $1 - p > .5$ regardless of the signal.

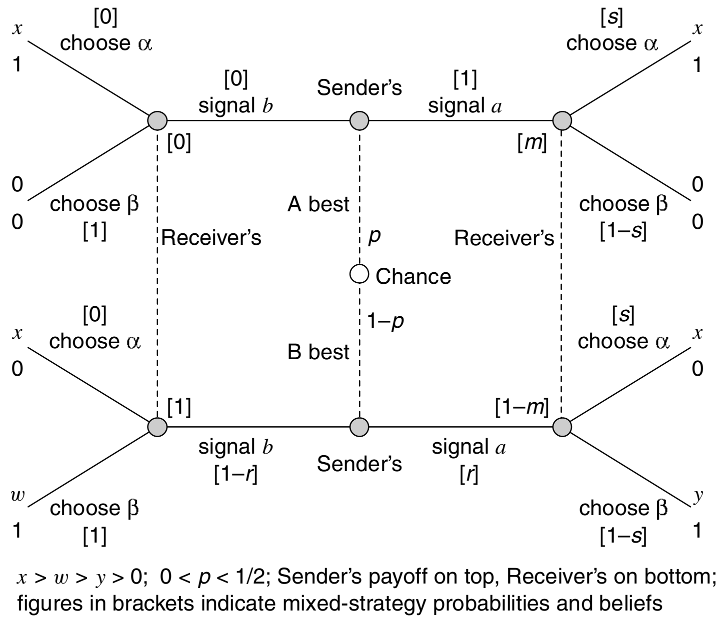


Figure 2 Semipooling Equilibrium in a Signaling Game

These are often referred to as "babbling" equilibria, since the Sender's signaling is meaningless.

Cooperative Game Theory

The term *cooperative game theory* should be understood to cover a variety of fairly distinct approaches, including axiomatic bargaining theory and spatial voting theory; but the central and original branch of the theory is that of games in *characteristic function* form or simply *coalition games*. A brief treatment will serve to illustrate the approach and its differences from and similarities to the noncooperative theory.

As the emphasis on coalitions might lead one to expect, some of the original applications of game theory to politics used cooperative game theory. The cooperative game formulation provided the basis for the calculation of *power indices* that, in a 1950s conceptual breakthrough, clarified the difference between voting weights and voting influence for small weighted-voting bodies and for voting bodies with complex veto rules, notably the United Nations Security Council. The earliest models of voting competition also took the form of cooperative games, being based on the specification of winning and blocking coalitions; this approach is still used productively, both in theoretical analyses and in the empirical study of multiparty election systems. A related early political science use of game theory was in the study of parliamentary government formation, where coalition patterns were analyzed using both party size and the restriction that coalitions include ideologically compatible partners.

The theory of coalition games originated with the attempt of John von Neumann and Oskar Morgenstern to theorize about zero-sum games with more than two players. In such games, a coalition of players could act in concert to assure themselves certain outcomes regardless of the efforts of the remaining players; the question is, "what coalitions will form to take advantage of their capabilities?" The original formulation assumed, contrary to our previous assumptions, that payoffs have meaning across players and can in fact be combined and voluntarily transferred from one player to another in a simple additive fashion. This class of coalition games is today called *transferable utility* (TU) games; in nontransferable utility (NTU) games, by contrast, one keeps separate account of the payoffs chosen by the coalition for each player.

Formally, let $N = \{1, 2, \ldots, n\}$ be a set of players. For every subset of players C, let $v(C)$ be the total payoff that those players can ensure receiving regardless of the actions of the remaining players; each nonempty C is a possible *coalition*, and v is the *characteristic function*; strictly speaking, v completely describes the game (since it implicitly includes a specification of the set of players). A valid characteristic function should be *superadditive*: For any two disjoint coalitions C and D, $v(C) + v(D) \leq v(C \cup D)$ since, acting together, the members of the two coalitions could do anything they could have done by acting separately, and perhaps more. Let $x = (x_1, x_2, \ldots, x_n)$ be the distribution of payoffs that ultimately results from the play of the game, after any transfers of utility to which members of coalitions have agreed. The aim of the characteristic function formulation is to predict the distribution of payoffs among the players, given the capabilities of the coalitions that they might enter into. Notice that the particular actions taken by individuals, acting alone or in concert, are completely abstracted away, and that it is implicitly assumed that a coalition can perfectly accomplish any desired commitment, cooperation, or coordination to achieve its potential total payoff. The definition of $v(C)$ as the most that C's members can ensure themselves embodies a pessimistic viewpoint that is really appropriate only where the underlying game is zero sum (so $v(N) = 0$), although that restriction is not always imposed when the coalition theories are applied.

The first, very weak, solution concept is the set of *imputations* of a game v. This is the set of payoff distributions satisfying two properties:

1. *Individual rationality*: For every player i, $x_i \geq v(\{i\})$, since the player has no reason to accept a payoff lower than what he or she could ensure himself or herself by acting alone.
2. *Efficiency*: $x_1 + x_2 + \cdots + x_n = v(N)$, since by definition the grand coalition can get any feasible total payoff, and there is no reason to settle for less.

The addition of one more condition defines the fundamental solution concept for coalition games, the *core*. Let x and y be two imputations; x is said to *dominate* y if there is some coalition C such that

(a) $v(C) \geq \Sigma_{i \in C} x_i$, that is, members of C have the power to impose x, and (b) $x_i \geq y_i$ for each $i \in C$, and strictly so for some $i \in C$, that is, each member of C is at least as well off, and some better off, under x than under y. The core of v then consists of the set of all undominated imputations. In many games of interest, the core is empty; consider, for example, three players who divide a dollar by majority rule. The Bondareva-Shapley theorem gives a *balance condition* ensuring the core will be nonempty. Alternative conditions on imputations yield other solution concepts, such as the *von Neumann-Morgenstern solution*, the *kernel*, and the *nucleolus*.

Randall Calvert
Washington University in St. Louis
St. Louis, Missouri, United States

See also Analytic Narratives: The Method; Bargaining; Cooperation; Coordination; Experiments, Laboratory; Prisoners' Dilemma; Rational Choice; Social Choice Theory; Spatial Models of Politics

Further Readings

Camerer, C. P. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games*. Cambridge: MIT Press.

McCarty, N., & Meirowitz, A. (2007). *Political game theory: An introduction*. Cambridge, UK: Cambridge University Press.

Morrow, J. D. (1994). *Game theory for political scientists*. Princeton, NJ: Princeton University Press.

Osborne, M. (2004). *An introduction to game theory*. New York: Oxford University Press.

Peleg, B., & Sudhölter, P. (2007). *Introduction to the theory of cooperative games* (2nd ed.). Berlin: Springer-Verlag.

Weibull, J. W. (1997). *Evolutionary game theory*. Cambridge: MIT Press.

Young, H. P. (1998). *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton, NJ: Princeton University Press.

Gender

Today, *gender*—a term that is also used in languages other than English—is used in the context of gender relations or as a social construction of sexuality. It is comparatively a new term, with *gender* originally being used until the 1960s only as a grammatical category to allocate nouns into different categories—feminine, masculine, and neutral. Since then, the term has been fully engrained in the social and cultural sciences discourse and has also established itself in the political discourse of different societies, as the concepts of gender mainstreaming and gender budgeting show. The expansion of the term and its widespread use can be seen as a sign that gender inequality is perceived in many areas of society and is examined as a social–organizational form. This entry discusses the origins of this concept, its relevance in social discourses, and its implications for contemporary political science.

Background

As opposed to the biological term *sex*, the term *gender* refers to the "social" or "psychological" gender of a person, where "sex" and "gender" should be seen as analytically independent of each other. Gender refers to a historically specific practice of social classification. It focuses on the special perceptions and patterns of interpretation by which the binary structure of "male" and "female" is viewed, dealt with, and institutionalized. In doing so, it deals with the symbolic order of sexes in a society, as well as with the self-attributions and identification of the individual, which focus on it but are not identical to this symbolism. In most societies, categorizing into male and female gender categories takes place on the basis of the shaping of primary and secondary gender characteristics. Disambiguations and the one-sidedness of "male" and "female" are associated with the two-gender classification, as well as with social relations of symmetry and asymmetry, equality and hierarchy, of inclusion and exclusion. These relationships are embedded in concrete social relations and structural connections. At the same time, the fact that classifications do not simply exist but constantly need to be reacquired points to their instability and dependency on having to be updated at a practical level. Based on the realization that social gender differentiation is not based on anthropological, biological, or physiological conditions but on the result of social classifications, it can be inferred that the hierarchical relationship of the sexes to one another is not an expression of a "natural" order that cannot be changed.

Stages in the Discourse

The sociological debate on inequality between women and men and the category of "gender" began around the middle of the 1960s and is closely linked to the so-called second wave of the women's movement emerging at that time in many Western countries. The studies, mainly compiled by female academics, first concentrated on unmasking science as "male oriented" and criticizing the lack of awareness of women's lives and active dealings in research and teaching. A close connection between the women's movement and feminist criticism in science was established, and it resulted in criticism of the discrimination against women in academia and society and of the one-sidedness of research that was mainly conducted by men and that often treated women as objects.

Feminist criticism of science highlighted the facts that cultural, political, and socially dominant gender images influenced the scientific theories of the time. Based on this, the much lauded scientific objectivity was often distorted by a male point of view. The primary aim in women's studies was, therefore, to make the ignored contributions of women in academia visible by coming to terms with and analyzing the social setting of the time and women's thoughts and actions. In doing so, women's studies were viewed at first as biased research that placed the focus on women and on the changes in their life situations.

This view was gradually broadened, however, and increasingly women's thinking and actions were analyzed in relation to men's with the intention of identifying differences and hierarchies. It was in this context that the discussion began about what sex and gender really mean. In gender and women's studies, gender is no longer treated as something natural but as a social or structural category with subsequent social roles. People are assigned or denied specific characteristics, attributes, and places in society on the basis of their sex. Sociological gender studies since the 1980s, therefore, speak of gender as a structural category. What is meant is that the separation and hierarchization of social spheres such as reproduction and production, and, consequently, the public and private realms are significant characteristics of modern, industrial societies. This separation and hierarchization are accompanied by an engendering of different spheres and social subdisciplines. Gender is regarded as a binary hierarchy that co-organizes the different levels of social organizations. Although study on the discrimination against women is frequently the point of departure for research, gender research now also addresses the social organization of gender relationships more broadly.

Based on the recognition that gender is a part and a result of a construction process in which "he" and "she" as individuals are involved, a sociological shift has occurred: Gender is no longer viewed as a structural category but also as a process category. Candace West and Don Zimmerman's concept of "doing gender," which was developed in 1987, differentiates itself from the current sex/gender distinction and the nature/culture concept transported with it. Doing gender understands gender attribution and gender identity as a continuous production process that is taking place in all human activities. This ensures that the gender discrepancy is generated by people continuously making themselves—and allowing themselves to be turned into—women and men in their everyday lives. It is a collective and individual process of construction through which, depending on the context and the situation, dual-gender classification is strategically implemented in practical social actions.

Similarly, a critical "men's studies" category is emerging that focuses on male disciplinary and normalization processes and calls into question the norms and realities of "hegemonic maleness." A central understanding within men's studies is the recognition that hegemonic maleness is responsible not only for the dominance of men over women but also for hierarchies among men. Beyond this, research on maleness has emphasized the fact that men are also subjected to one-dimensional gender assignations that are formative and do not at all represent the "general human being." An important difference from women's studies continues to be that the experience of injustice and discrimination has not formed the basis of research and the relationship between research and practice in the context of a "political" or partisan intervention—as was central in the early days of women's studies—has therefore had very little significance.

Gender studies further broadened in scope in the 1990s. The category of gender, which had until then been viewed as a "natural" classification, even if culturally and socially formed, was questioned. *Gender Trouble*, a book by the American philosopher Judith Butler published in 1990, evoked a comprehensive discussion on the identity-giving category of "woman." What had widely served as an unquestionable basic assumption in women's studies—the natural existence of two exclusive genders—increasingly became an object of academic examination. The binary distinction between two—and only two—genders is now frequently understood as an ideological construct of a discursive power system. Gender affiliation of persons and the two-sex model as a social classification and functional differentiation principle do not refer to the law of nature but are the result of social constructive processes.

This construction of gender cannot, however, be viewed as a neutral form of determination; rather, it is already part of the social determination of gender relations and therein also of the social inequality between men and women. The factitiousness of gender not only affects gender, however, but also applies to the biological category of sex. As the latter is also not free of a sociocultural construct, the "prediscursivity" of sex and the norms and regulations tied to biological sex must be debated.

Gay theory also focused the spotlight on the connection between gender and heteronormativity, or the social norm of heterosexuality. Academic studies continue to take up heterosexuality as a power configuration, which gender studies has barely researched and has not acknowledged as hetero-normatively composed. Normative reference points in women's and gender studies are put up for discussion with the transformation and broadening of knowledge on the category of gender as a binary construct.

In the beginning, gender studies were oriented—as was the women's movement—toward the issue of how equality between women and men could be attained in all areas and which barriers are an obstacle to this goal at the political, legal, symbolic, and other levels. At a practical political level, equality was seen as nondiscriminatory, but this view raised the issue that equality leads to a cementing or even strengthening of inequality according to the different points of departure and socially created differences. In addition, there was a criticism that the concept of equality was targeted toward aligning the position of women—their social and individual possibilities—with those of men. The measure of equality is thus "the man" and the options, values, and concepts of work, life, and politics ascribed to him.

As a patriarchal devaluation of the female is latently connected to this view of equality, the difference perspective articulated the suggestion that women should be treated differently from men. Connected to this is the challenge to go beyond stereotypical attributions to reflect on the ways in which women are different or unique. Furthermore, the characteristics, lifestyles, and so on of women should not be viewed as deficient but as parts of one and the same concrete whole. Associated with this, however, is a fear of a codification of typical female characteristics and gender roles that tend to become gridlocked. The discussion on equality or difference has become a never-ending story; nonetheless, it is clear that the deconstruction of gender has led to a shift from the selection of one of these two positions to a concern with analyzing the connections between equality and difference. For some years now, the perspective has broadened, and in addition to the gender imbalance, differences among women are also increasingly being taken into consideration—especially class-specific or ethnic differences that resulted from American theory-building discourses and Black feminism, as well as the debate surrounding women (co)perpetrators during National Socialism.

As all individuals not only belong to *a* group but always belong to multiple groups at the same time, how the category of gender interacts with other categories of inequality is analyzed below. Both legally and politically, this is codified in the concept of multiple discrimination. At a theoretical level, gender studies subsume this phenomenon as a crossing over between different categories of inequality under the term of *intersectionality*, a term that the American legal specialist Kimberlé Crenshaw coined in 1987 in the context of Black feminism. The debate surrounding intersectionality and the interdependencies among race, class, and gender has so far been mainly theory driven and is only gradually being enhanced by empirical studies. In general, it is argued that the question of whether sex or another category has a larger influence on discrimination, which has been raised in the discussion on gender as a "master category" from time to time, should remain open in the analysis.

Gender in Political Science

In comparison with other social science disciplines, political science was relatively late in focusing on gender as a category. In the beginning, solely the field of "women and politics" was the object of study. Within this context, the question of the political representation of women and how it was reflected in political issues was especially interesting. The feminist perspective first took up this analytical perspective and questioned the reasons behind the formal and informal exclusion of women from politics.

A reconceptualization and expansion of the research field to include "gender and politics," however, was initiated relatively quickly. In this context, maleness/masculinity as a measure and norm is inscribed in all political processes in a more or less subtle manner. In most countries, a quantitative overrepresentation of men in politics exists that is oriented on the biological sex. Politics is not, however, male because almost all political decision-making bodies are dominated by men, but because male interests are systematically presented as hegemonic, representing and giving preferential treatment to interests that most closely correspond to the everyday lives of this social group. In addition, there is a substantive masculinity: as Joan Acker highlighted with the term *gendered organizations*, institutionalized male patterns are integral components of political organizations that influence political rules, values, norms, and structures.

Thus, gender-critical discourse within political science generally takes place at two levels. First, it occurs at the sociopolitical level and expounds on the limited number of women in the discipline. The underrepresentation of women results in the fact that women receive less attention as actors and, at the same time, the androcentric scholarly conception of women, sex, and sexuality is maintained. The other level focuses primarily on epistemological and conceptual questions and delivers a critical inspection of political science paradigms and terms at a theoretical analytical level with the goal of making the gender subtext visible. Based on this deconstruction of apparently gender-neutral normative premises, terms, and research interests in mainstream political science, the political was broadened and reformulated to include a gendered view. Part and parcel of this was a critical revision of its core areas and subdisciplines, among others the fundamentals of political scientific thinking. Among these, especially, are the fundamental categories of the public sphere, privacy, and politics.

The Public Versus the Private Sphere

The central point of departure for a gender-critical political science is the challenge of the separation and polarization between the public political and the private nonpolitical spheres. Every debate concerning this distinction deals with the problem of participation at its very core or the possibilities and conditions of women's presence in politics. In the history of political ideologies, "women, together with other groups such as slaves, servants (to which a large number of women belonged) and landless men, were described as incapable of leading a moral life or able to participate in the justice system" (Seyla Benhabib & Linda Nicholson, 1987, p. 514). The differences and inequalities between the genders were substantiated on four assumptions: First, the natural, anatomical, and physiological differences were used to justify inequality by claiming that the ability to give birth as well as other apparent characteristics of women were decisive criteria for participating in political life. Second, it was assumed that men are both physically and intellectually superior to women, which brought with it a hierarchy of naturalized differences that was also reflected in the significance of the spheres. Third, the so-called female sphere is irrelevant for political life. This is accompanied by the fourth point—the assumption that the domestic sphere, although it takes place outside the political arena, is nonetheless controlled by politics and is inferior to it (Benhabib & Nicholson, 1987, p. 515). In sum, this means that male reason and rational conduct are attributed to the public sphere, while the private sphere is seen as being characterized by irrational, unpredictable patterns.

Criticism is targeted at the double separation of the public and the private, which has traditionally influenced theory formation in political science. By distinguishing state and (private) economy on the one hand, and economy and (private) households on the other, gender inequalities are left out of sight in many respects. The gender-specific division of labor that is constitutive for the (private) economy (paid/unpaid) eludes state and political regulation, such as gender-hierarchical appraisal of labor (well paid/badly paid), which essentially represents a continuation of this labor division in a semipublic space in its differentiation between women's and men's activities. The entire vocabulary used in political science is built on this dichotomy. Feminist welfare state research especially points to the fact that gender-specific labor division is not only a constitutive element of the (welfare) state—as Carole Pateman (1988) showed with the "sexual contract" approach—but that politics is also gendered and supports gendered processes: "In a society in which there is a gendered division of labour there is almost no area of policy in which women and men are not differently affected" (Joni Lovenduski, 1992, p. 610). The discourse on the separation of public and private highlights the fact that a close connection exists between the two spheres, and that, at the same time, these are placed in a hierarchy (e.g., important–unimportant, inclusion–exclusion) that is constructed and has an ideological basis.

Broadening of the Definition of Politics

The division between the public and the private sphere has also become engrained in the field of politics and the political. The classic definition of politics views politics (*politikos*) as the administration of the affairs of the *polis*, the community. In contrast, the areas of housework, household administration, the house (*oikos*), and the family, or in other words those areas in which women were traditionally active, were not seen as political. Even if this classic political understanding in political science does not dominate (anymore), and many different political definitions exist, the political is commonly positioned in the public sphere. Feminist political science, therefore, calls for a widening of a public and institutionalized definition of politics that only sees gender relationships and gender relations as un- or prepolitical but not as a political expression or object.

The American political scientist and philosopher Nancy Fraser, for example, distinguishes between an official political and a discursive political definition. The former views a matter as political if it is represented in official government institutions or, in other words, in parliaments, administrations, and other bodies. As Fraser (1989) notes, "the discursive political definition comprises everything that is disputed beyond a range of different discursive arenas and among a range of different discourse publics" (p. 166). Between the political definitions is a close connection, as an issue usually only turns into an object of legitimate state intervention once it has been discussed among the broader public. The border between both attributions is not defined but is an object of conflict and changes within social discourses.

The expansion of political contents and spaces is accompanied by criticism against forms, norms, processes, and practices linked to a governmental political definition. This would lead to the fact that women not only as a group of people but also concerning women's interests (even if these are not necessarily homogeneous interests) would insufficiently be represented or not be represented at all.

A further important aspect in the discourse surrounding the definition of politics concerns the question of how gender relations and gender relationships can be viewed as political objects.

When politics govern social relationships, they also embody identities and interests and the relationships between men and women, as well as between different groups of women or men. If this understanding is pursued, the field of activity of political science is broadened to include a debate and analysis of politics as the regulation of bipolar hierarchical gender relationships. Birgit Sauer (2005) outlines the meaning of gender categories for political science as follows:

> As a structural category in politics and political science, gender is constituted at the subjective individual and objective levels—a culturally formed, socially shaped relationship that is loaded with purpose and meaning, that possesses structural and institution-building potency. (p. 381)

Gender is to be understood as a structural and structuring category. This means that gender, on the one hand, is a political institution and that, on the other hand, institutions "have" a gender and reproduce genders. The following five levels are relevant to the analysis of gender in political science: first, the level of political cultures and symbols; second, the level of state organizations and rules (polity); third, the policy level; fourth, the level of political processes and their actors (politics); and fifth, the level of political subjects.

Status Quo

Nowadays, a gendered view of the discipline can be found in nearly all areas of the field, such as international relations, political theory, political systems, or policy research. Similar to this profiling, a partial institutionalization of feminist or gendered political science as an area of research and as a teaching subject has taken place. Nevertheless, a problem that Virgina Sapiro (1995) identified almost 15 years ago continues to be true:

> Despite the size of the gender politics community, and despite the fact that most political science departments now offer at least one course in this area, gender politics has not been fully integrated into political science; its theories, questions, and conclusions have not been "mainstreamed" to any significant degree. (p. 292)

This means that gender relations beyond the feminist and gender-oriented branch of the discipline are not, or are barely, conceptualized as social relations that structure politics.

Alexandra Scheele
University of Potsdam
Potsdam, Germany

See also Feminism; Feminist Movements; Feminist Theory in International Relations

Further Readings

Benhabib, S., & Nicholson, L. (1987). Political philosophy and the question of women. In M. Lotsch & W. Preller (Eds.), *Piper's handbook of political ideas* (Vol. 5, pp. 513–562). Munich, Germany: Piper.

Carroll, S. J., & Zerilli, K. M. G. (1993). Feminist challenges to political sciences. In A. W. Finifter (Ed.), *Political science: The state of the discipline II* (pp. 55–76). Washington, DC: American Political Science Association.

Fraser, N. (1989). *Unruly practices: Power, discourse, and gender in contemporary social theory.* Cambridge, UK: Polity Press.

Lovenduski, J. (1992). Gender and politics. In M. Hawkesworth & M. Kogan (Eds.), *Encyclopedia of government and politics* (Vol. 1, pp. 603–615). London: Routledge.

Pateman, C. (1988). *The sexual contract.* Stanford, CA: Stanford University Press.

Sapiro, V. (1995). Feminist studies and political science—and vice versa. In D. Stanton & A. Stewart (Eds.), *Feminism in the academy* (pp. 291–310). Ann Arbor: University of Michigan Press.

Sauer, B. (2005). Limitation and transgression of the political gender studies in political sciences. In H. Bußmann & R. Hof (Eds.), *Genus: Geschlechterforschung/gender studies in den Kultur- und Sozialwissenschaften* [Genus: Gender studies in cultural studies and social sciences] (pp. 366–401). Stuttgart, Germany: A. Kröner Verlag.

Generalized Least Squares

See Weighted Least Squares

Genocide

While the term *genocide* is new, the phenomenon is old. Genocide has a long history, going back to at least the 12th century BCE and continuing into the present. The 20th century can be considered an "age of genocide," with numerous examples of genocide throughout the period claiming as many as 70 million lives (some scholars would say many more than that). After the Holocaust, the watchword was "Never again," yet since 1945 the lives claimed by genocide exceed those killed in international and civil wars combined. But with so many genocides, different situations, motives, and victim groups, it has been difficult to define *genocide* in ways that everyone would accept. Moreover, the term is widely used in a rhetorical fashion to attract attention and to further political and social demands.

Nevertheless, the definition of genocide that is contained in the United Nations (UN) Convention on the Prevention and Punishment of Genocide, though widely criticized, is the one that prevails among scholars. To complicate matters further, the concept of genocide could be described, following Ludwig Wittgenstein, as a concept with "blurred edges." Here, the family resemblances and differences between genocide and other concepts, such as ethnic cleansing, must be taken into account. There is still a core meaning: Genocide consists of acts that intentionally threaten the existence of human groups. Yet this statement itself is incomplete: What acts? What level of threat? What is a group? Who defines it? And how is intentionality determined?

Before the term *genocide* existed, mass killings were referred to as *massacres*, *race extermination*, and so on. Perpetrators had very different terms: *victory*, *self-defense*, *purification of society*. It was not until 1944 that the term *genocide* was created by the Polish Jewish lawyer Raphael Lemkin in his book *Axis Rule in Occupied Europe*. The term was based on the Greek *genos* (race, group) with the Latin *cide* (killing). Lemkin used the term to describe not so much the destruction of individuals but "the destruction of a nation or ethnic group." Political and social groups, such as classes, were not included in his definition and nor was the partial destruction of a group. He stated, "Genocide is a coordinated plan of different actions aiming at the destruction of the essential foundations of the life of national groups, with the aim of annihilating the groups completely" (p. 79). For him, the core of the crime was the destruction of groups, not individuals; he believed that groups were the bearers of civilization and culture, of identity, of diversity, and thus in special need of protection. Pieter Drost (1959) countered with a definition that focused not on the destruction of groups but of individuals: "Genocide is the deliberate destruction of physical life of individual human beings by reason of their membership of any human collectivity as such" (p. 125).

Lemkin pressed the UN to declare genocide a crime against international law. From 1946 to 1948, the UN debated the wording of what would become, by unanimous vote in December 1948, the UN Convention on the Prevention and Punishment of Genocide. With sufficient ratifications by member states, the Convention entered into force in early 1951. But despite the occurrence of numerous examples of genocide in the years following, not a single case of genocide was tried under international law until the opening years of the 21st century. The International Tribunals for the former Yugoslavia and Rwanda have now both charged and convicted several individuals for genocide. One problem with this, however, is that a vast, state-organized process, with deep societal involvement, is treated as if it were simply a trial for murder.

The Convention makes genocide and related crimes, such as incitement to genocide, crimes under international law, "whether committed in time of peace or in time of war." But the heart of the Convention is found in Article II, where the crime of genocide is defined and where many of the problems with the concept lie:

> In the present Convention, genocide means any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group, as such:
>
> (a) Killing members of the group;
>
> (b) Causing serious bodily or mental harm to members of the group;
>
> (c) Deliberately inflicting on the group conditions of life calculated to bring about its physical destruction in whole or in part;
>
> (d) Imposing measures intended to prevent births within the group;
>
> (e) Forcibly transferring children of the group to another group.

Many important issues arise from this conception of genocide. First, not all groups are protected; political and social groups, such as classes, lie outside the boundaries of the definition. The reasons for this are that (a) only groups would be covered that were "stable and permanent" and (b) many countries wanted to keep open their right to suppress by violence any political or social challenge to the authority of the ruling power. Second, not only must there be intent to destroy a group, but there must also be "specific intent," that is, the intent to destroy the group must be because of its particular characteristics (national, ethnical, racial, or religious). Destroying such a group for, say, economic reasons, would not count as genocide. Third, the question of destroying a group "in part" has raised issues about how large a part and whether it is a matter of numbers or a percentage of the overall group that counts. In general, these issues have now been dealt with by saying a "substantial part," leaving it to courts to decide just where to draw the boundaries. Some legal scholars, however, would add that "in part" could mean a "significant part," in some functional sense—for example, culling the leadership elements of the group.

A fourth issue concerns the problem of what constitutes a group. For example, the Tutsis in Rwanda were neither a national nor an ethnic group, but they were perceived by their killers to be members of a group. In the UN Rwandan tribunal, the court decided that groups and membership in them are defined by the perceptions of the perpetrator. Finally, the acts that are listed as constituting genocide are physical, biological, and cultural may not matter, but of the five acts, any one of which constitutes genocide, only one of them actually involves killing, though all aim at the destruction of the group. The UN definition focuses on intent, not on results. That genocide could occur with few or even no deaths is a surprising conceptual outcome of the UN's approach and contrary to both scholarly analysis and public understanding of what is meant by genocide.

There are other concepts that bear a family resemblance to genocide but also can be distinguished from it—for example, war. Many genocides occur within the context of war, but apart from ancient warfare, which was synonymous with genocide, war traditionally was a battle between armed combatants not between perpetrators and largely unarmed civilians. In the 20th century, war has often been total war, and waged against civilians as well as soldiers, but not with the idea of eliminating the "enemy" as a group. Genocide, on the other hand, is a species of total war: It is aimed at a particular group, most of whom are unarmed, and the intent is not the subduing of the group but its elimination, if not in whole, in substantial part.

Massacre also can be part of how genocide is carried out, but more often it is geared toward revenge or repression, not annihilation. It tends also to be episodic and isolated rather than sustained and pervasive as in genocide.

The term *ethnic cleansing*, on the other hand, has gained so much traction since the war and genocide within Bosnia that it has almost displaced the idea of genocide or, alternatively, is treated as another term for genocide. In practice, ethnic cleansing is often part of the genocidal process; the term is also a euphemism used by the perpetrator to disguise what is actually taking place. But in theory, ethnic cleansing means using violence, rape, terror, and forced deportation to drive out or remove a particular group from a common territory. The emphasis is on creating a homogeneous society through fear and expulsion, not as in genocide on annihilating the group.

There are legal and institutional developments that also relate to the status of genocide within international law. In 1998, the Statute of Rome created an International Criminal Court, with the following crimes within its jurisdiction: genocide, crimes against humanity, war crimes, and the crime of aggression. The very term *crimes against humanity* arose out of the Young Turk genocide of the Armenians, when the Allies in May 1915 warned the Ottoman rulers that they would be held accountable and tried for "crimes against humanity and civilization." In the statute, "crime against humanity" is defined in terms of a number of particular acts, including murder, extermination, and deportation or forcible transfer of population but more widely, torture, apartheid, and "other inhumane acts of a similar character intentionally causing great suffering, or serious injury to body or to mental or physical health." Any of these acts "when committed as part of a widespread or systematic attack directed against any civilian population" constitutes a crime against humanity. The key category here is "civilian population," not particular groups as indicated in the Genocide Convention; and the question of intent is different from that required under the Convention: Here, no specific intent is required.

Though one can differentiate "genocide" and "crimes against humanity" in important respects, the two concepts overlap in many ways. But it is much easier to prosecute actions as crimes against humanity than as genocide; the former also have a wider reach, not being restricted to crimes against particular groups named in the Convention. Some international legal scholars suggest that with the increased emphasis on crimes against humanity, the international community could dispense with the Genocide Convention. Prosecution of crimes against humanity would thus provide alternatives to the conceptual confusions and limitations of the Convention.

Roger W. Smith
College of William and Mary
Williamsburg, Virginia, United States

See also International Law; Violence; War and Peace

Further Readings

Chalk, F., & Jonassohn, K. (1990). *The history and sociology of genocide*. New Haven, CT: Yale University Press.

Drost, P. (1959). *The crime of state: Vol. 2. Genocide. United Nations Legislation on International Criminal Law*. Leiden, Netherlands: A. W. Sijthoff.

Lemkin, R. (1944). *Axis rule in occupied Europe*. Washington, DC: Carnegie Endowment for International Peace.

Schabas, W. (2000). *Genocide in international law*. Cambridge, UK: Cambridge University Press.

Smith, R. (1987). Human destructiveness and politics: The twentieth century as an age of genocide. In I. Wallimann & M. Dobkowski (Eds.), *Genocide and the modern age* (pp. 21–39). Westport, CT: Greenwood Press.

Straus, S. (2001). Contested meanings and conflicting imperatives: A conceptual analysis of genocide. *Journal of Genocide Research, 3*(3), 349–375.

United Nations. (1948). *United Nations Convention on the Prevention and Punishment of Genocide*. Geneva: Author. Retrieved November 19, 2010, from http://www.hrweb.org/legal/genocide.html

United Nations. (1998). *Rome statute of the International Criminal Court*. Geneva: Author. Retrieved November 19, 2010, from http://untreaty.un.org/cod/icc/statute/romefra.htm

GEOPOLITICS

Geopolitics denotes, generically, strategic power struggles among states over the political control of territories and resources. More specifically, it suggests that politics is primarily defined by geographical location, territorial expansion, and interstate competition over finite spaces. The late-19th-century interimperial rivalries provided the historical context for the rise of geopolitical thought across the core countries of the imperial zone, rearticulating the relation between economic expansion and territorial control. The final division of the last extra-European vacant spaces generated an acute awareness of the closure of absolute space in a new "planetary" age. Inter-imperial relations came to be regarded as intensified zero-sum conflicts over the redivision of an occupied planet. This closed spatial horizon prompted a reconceptualization of the state as a territorial phenomenon in space, locked into a permanent struggle for survival. The revalorization and politicization of geography and its absorption into a reconceptualized science of politics forged the new field of political geography—geopolitics' direct precursor. Today, particularly with increasing globalization, geopolitics is more and more challenged and questioned: Territory has lost its classical relevance, as transnational relationships imply a new vision of the global world, which is less compatible with geopolitics. Critical geography points out the ideological and political background of geopolitics.

The concept of geopolitics was invented by the Swedish political scientist Johan Rudolf Kjellén (1864–1922) and developed in *The State as an Organism* (1917)—a critique of the prevailing legal positivism in constitutional and international law. The radicalization of political geography into *Geopolitik* as a specifically German theory of international politics was prosecuted—most notably by Karl Haushofer (1869–1946) and his school and, from a different angle, by Carl Schmitt (1888–1985)—during the revisionist conflicts over the Versailles Settlement in Weimar Germany. Geopolitics had a lasting, though not exclusive, impact on Nazi foreign policy. Intellectually, it positioned itself as a counternarrative against Marxist theories of imperialism (Rosa Luxemburg, Nikolai Bukharin, Vladimir Lenin) and against the Wilsonian liberal internationalism, represented by Isaiah Bowman (1878–1950). Basic elements of geopolitical thought infused, if in a semantically altered and politically sanitized form, the American Cold War discourse of neorealism in a new binary world-political geography. The proliferating but often indiscriminate contemporary invocations of the term *geopolitics* reflected the shifting post–Cold War strategic geography—globalization, multipolarity, hegemony, unipolarity, empire—and a renewed urgency of resource politics in times of natural resource depletion and environmental degradation. Simultaneously, the new approaches of critical and Marxist geopolitics sought to recover, respectively, the cartographic discourses that shape the construction of geopolitics and the efficacy of geopolitical conflicts that mediate European and worldwide socioeconomic and political developments.

Friedrich Ratzel (1844–1904), a professor of political geography and the cofounder of the *Alldeutscher Verband* (Pan-German League), was cross-nationally the most influential figure in the development of the geopolitical tradition. Applying ideas from biology, evolutionism, and human geography, he contended that state behavior—the expansion and contraction of states in space—is primarily determined by geographical properties and geostrategic location. Space, as a political category, is no longer conceived as a geometric, neutral, and empty expanse (the absolute space of nonhuman geography) but as a concretely ordered and constructed territory (the historical-relational space of political geography). Fusing an organic-biological conception of the state with principles of Darwinian natural selection in interstate relations, Ratzel claimed that states come to life, grow, and die in the struggle for space. Borders are temporary phenomena in a pulsating geopolitical environment. Preoccupied by the quest for a "scientific" legitimation of Wilhelmine Germany's expansionist policies, his *Politische Geographie* (1897) suggested that states enjoy a natural right to an adequate *Lebensraum* (living space) grounded in variations between soil fertility and population growth. Population "density pressures," established by a calculus between state territory and demography, are formalized in the "laws of the spatial growth of states." Differences arise in the prewar era between ethnocentric and geopolitical conceptions of states and the goals and limits of their foreign policies. While the ethnocentric notion defines nations as cultural-linguistic units (*Volksnation*), leading to a self-limiting territorial correspondence between a nation's area of settlement and the scale of state territory, the latter conception (*Staatsnation*) prioritizes territorial aggrandizement over ethnic-racial homogeneity, although it may also involve an active policy of ethnic settlement (Germanicization). *Volksnation* and *Staatsnation* are not synonymous and constitute rival points of reference. This tension between "race" and "space" resurfaced later in differences between the original program of German geopolitics and Adolf Hitler's racial ideology.

Contemporaneously, Alfred T. Mahan's (1840–1914) anti-isolationist navalism and Halford Mackinder's (1861–1947) heartland theory formed the core of political geography in the United States and Britain. Admiral Mahan's *The Influence of Sea Power on History, 1660–1783* (1890) identified the fleet-based domination of seaborne commerce as the decisive factor in world politics, exemplified by Britain's early modern oceanic hegemony. In an age of pre-aerial warfare and limited mechanization of terrestrial military technology, a major battle fleet secures a mutually reinforcing and self-sustaining relation between naval domination and commercial supremacy—the geostrategic axis of history. Mackinder drew the opposite conclusion in his lecture *The Geographical Pivot of History* (1904). Naval power, after unifying the world into a closed political system during the Columbian age (1500–1900), was superseded during the post-Columbian age by land power. According to Mackinder's trizonal global geography, a geostrategically immune inner-Eurasian heartland—the pivot of history—is surrounded by an inner crescent (the European periphery, Near/Middle East, India, China), enclosed, in turn, by an outer crescent (England, the Americas, Africa, Australia, Oceania, Japan). The contemporary global balance between land and sea power was decisively altered by the infrastructural penetration of the pivot area (Central Asia, Russia) through the construction of transcontinental railway networks. The future belonged to a German–Russian alliance, strategically invulnerable in its territorial heartland against sea power. Only a grand coalition of sea powers could prevent Eurasian world domination. Both conceptions had to be situated in the transition from cosmopolitan free trade imperialism to protectionist neomercantilism.

During the passage from the 19th-century European pentarchy to the new constellation of world politics, flanked by the rise of the United States and the former USSR, political geography turned into German geopolitics. Kjellén's antilegalistic *The State as an Organism* contrasted the norm-oriented concept of the state as a legal subject, most notably represented in Weimar Germany by Hans Kelsen, with an "empirical viewpoint" of the state as a geographic organism commanding space. Foreign policy was elevated to an existential condition. Geopolitics was redefined as an auxiliary science of statecraft in the international struggle for survival. Ratzel's and Kjellén's thoughts received a wide and enthusiastic reception in Weimar Germany during the revanchist intellectual struggles against the Versailles *Diktat* and as a key legitimation for national-socialist *Großraumpolitik*. The rise of the geopolitical tradition in the 1920s was marked by the foundation of the Institute of Geopolitics in Munich (1922), the *Geopolitische Seminar* at the *Deutsche Hochschule für Politik* (1924) in Berlin, and the launch of the high-circulation *Journal of Geopolitics* (1924). Haushofer, who was professor of geography at the University of Munich (1919–1939), director of the Institute of Geopolitics, president of the German Academy (1934–1945), and a close confidant of Rudolf Hess, emerged as the most influential figure in the German discourse. Its intellectual core was premised on

1. an organic-biological conception of the state and a Darwinian view of interstate rivalry;
2. a critique of Western, liberal, and "mechanistic" conceptions of state and society, contrasted with the *völkisch* "ideas of 1914" that recast states as expressions of culturally defined and homogeneous peoples, organically rooted in specific territories;
3. a revaluation of agriculture, mystified as "chthonic" and "organic," coupled with a neo-Malthusian understanding of population growth, leading to demands of an autarchic *Lebensraum*;
4. the rejection of international law (League of Nations) and its replacement by a new *Großraumordnung* (Grand Spatial Order);
5. a geostrategic determinism premised on the dualism between land power and sea power, informing the programmatic policy prescriptions for a Eurasian power bloc under German leadership;
6. the promotion of geography as the key science of statecraft and the dissemination of "suggestive cartography" for the collective preparation of the nation in its struggle for survival; and
7. the combination and political radicalization of these elements in Anglophobia and anti-Semitism.

Geopolitics matured from 1933 onward into the official German science of the state. Scientific pretense and normative orientation blended into a highly politicized literature. Attacking the "liberal" system of European *Kleinstaaten* (ministates), the juridicization of international politics and the fixing of borders, the geopoliticians invoked during the 1930s the "law of growing spaces." Buoyed by the first Nazi foreign policy successes, the new spatial order envisaged no longer sovereign nation-states as the units of the international system—the defunct Westphalian System—but a plurality of *Großräume* (Pan-regions). Schmitt, although institutionally unconnected to the geopolitical school, emerged as the sharpest critic of the League of Nations project and advocate of a new *Nomos of the Earth* (1950). He argued that all conceptions of international law are ultimately grounded in metajuridical foundations—the power projections and land appropriations that underlie state formations and expansion, exemplified in relation to the Spanish and English conquests of the Americas. Anglo-American international law was based on a structural correspondence tying the postwar promotion of liberal-constitutional *Kleinstaaten* on the European continent and beyond to their capitalist penetration by private economic forces and integration into the world market. Their incorporation into the League obliterated their political essence—the right to define an enemy and to wage war—due to the turn toward a discriminatory concept of war. The apolitical "spaceless universalism" of universal law—the flattening of differentially organized polities and their submission to common legal principles—sanctioned a constitutive dualism between the proliferation of liberal-constitutional states and the expansion of a borderless private world market. This strategy represented a concrete political project of global domination—the rationalization of global space driven by a nonterritorial capitalist imperialism for American *Lebensraum*.

The architect behind this project was Bowman, director of the American Geographical Society (1915–1935), key adviser to Woodrow Wilson at Versailles and, later, to Franklin D. Roosevelt. Wilson's moralistic liberal internationalism was commensurate with the American national interest in an attempt to move international order "beyond geography." Simultaneously, Schmitt contended that U.S. hemispheric policy toward the Pacific, Middle, and South America (Monroe Doctrine, 1823) and the British Empire constituted legal precedents for his multi-*Großräume* vision. The Monroe Doctrine manifested Anglo-American double standards in international law and politics. Against this background, Schmitt set out the historico-legal argument for the pluralization and regionalization of diverse, coexisting, and mutually exclusive legal spheres. *Großraumwirtschaft* (Grand Economic Order) and *Großraum*—not a biological, but a political-juridical category—merged into a supranational hegemonic order in Central and Eastern Europe under German leadership, located between the Soviet bloc and the Western powers' spheres of interest. Schmitt remained vague on the inner structure of Pan-regions and silent on inter-*Großraum* relations, excepting the principle of nonintervention for foreign powers. Similarly, Haushofer pleaded with reference to Mackinder's heartland theory for the construction of a continental bloc, comprising Germany, the former USSR, and Japan. This conception informed the German-Japanese Anticomintern Pact (1936), the German-Soviet Non-Aggression Pact (1939), and the division of Poland. The contradiction between Haushofer's anti-Bolshevism and his demands for an Eurasian bloc remains irresolvable. During the late 1930s, the term *geopolitics* became recharged with *völkisch* terms and was finally subordinated to racial policies.

Benno Teschke
University of Sussex
Brighton, United Kingdom

See also Anarchy; Balance of Power; Biology and Politics; Capitalism; Fascism; Imperialism; International Law; International Relations, Theory; International System; Marxism; Power and International Politics; Realism in International Relations; Sovereignty; State; Territory; War and Peace; Westphalian Ideal State

Further Readings

Bowman, I. (1921). *The new world: Problems in political geography*. New York: World Book Company.
Ratzel, F. (1897). *Politische geographie* [Political geography]. Munich, Germany: Oldenbourg.
Schmitt, C. (2003). *The nomos of the earth in the international law of the ius publicum Europaeum*. New York: Telos Press. (Original work published 1950)
Smith, N. (2003). *American empire: Roosevelt's geographer and the prelude to globalization*. Berkeley: University of California Press.
Teschke, B. (2003). *The myth of 1648: Class, geopolitics and the making of modern international relations*. London: Verso.

Globalization

Globalization refers to the integration of separate nations, regions, or even individuals into a wider global system. It is characterized by increasing the linkages and connections between peoples and countries and by the growing knowledge of these interactions. This integration process can affect the economy, polity, society, or culture; that is, the definition of the process is very general and can refer to economic, political, social, or cultural integration. The question of who is being integrated is also left open, although in political science it is mostly the countries themselves that are at issue. The way in which such integration operates is also variable. The most common form of globalization in the literature involves the economic integration of national economies into a wider global one, usually through international trade, capital flows, labor movements, and global production networks. For some authors, this means the integration of countries into a global capitalist system—that is, one based on private property, limited government intervention, and the use of markets to allocate economic value. Some term this *neoliberal globalization*. Capitalism, international capitalist groups, and neoliberal practices all spread throughout the world as globalization proceeds. Whether this is good or bad for countries and individuals greatly depends on a scholar's perspective. Some associate globalization with the homogenization of distinct groups; one example is the loss of many local languages and the increasing diffusion of the use of English. Others associate it with the domination of American (or Anglo-American) values, beliefs, and practices. The global turn to liberal democracy is often cited as an example. Globalization is not complete at this point. Countries have different levels of integration into the world system. The extent to which a truly global system—whether economic, political, or social—exists is a point of debate. The causes and consequences of globalization are also debated.

Globalization has both secular and cyclical aspects. For some, the roots of globalization are in technological changes, such as the lowering of the costs of communications and transportation, which make the movement of people, ideas, goods, services, and capital faster and less costly. These changes are irreversible and often increase over time. While countries and their governments may try to put up barriers against such changes, it is very difficult for them to successfully fend off the pressures of globalization. In this view, globalization has been increasing steadily over time as technological change makes the world ever smaller. Sometimes technological change is treated as exogenous—that is, it has sources other than countries' policies or global integration. Other times, globalization itself, especially via trade of capital flows, is seen as having an impact on technology; it usually is conceived of as helping induce innovation and speeding its transfer around the world. Globalization and technological change may then go hand in hand, each encouraging the other.

On the other hand, globalization has been viewed as cyclical; high levels of global integration have been associated with two distinct periods in recent history. In the mid- to late 19th century, a global economy emerged as the European powers, especially the British, forged an open economy using both military power and economic policy. The development of colonies around the globe and the extension of European trade, capital, values, and political power undergirded this period of globalization. In the late 19th century, the world economy attained levels of openness that had never been seen before. In some areas such as the movement of labor, such levels of integration have never been achieved since. This open economy collapsed in the early 20th century with the two World Wars and the Great Depression. From 1914 to 1945, globalization was in retreat. Protectionism became the dominant trade policy for many countries; trade and currency blocs formed regionally, often around one powerful country, and military conflict was prevalent. Nationalism, isolationism, and regionalism marked this 30-year period. Globalization thus declined during this period.

However, after 1945 the United States picked up the mantle from the British and began to forge a world economy through the creation of international institutions, an open economic policy, and an internationalist foreign policy. The global system, however, was divided into at least two blocs with the United States on one side and the former Soviet Union on the other; hence, globalization would always be limited in this Cold War system. The United States pressed ahead with the organization of an open, capitalist economy guided by principles embodied in various international institutions. It worked to encourage democratization in certain regions and then to include democracies in this integration process. The creation of the General Agreement on Tariffs and Trade (GATT) and the subsequent development of the European Community started a process of trade liberalization among developed countries. The Bretton Woods monetary system supported by the International Monetary Fund (IMF) and the World Bank helped countries establish currency convertibility and a fairly stable world monetary system based on the dollar. This stability provided the conditions for the development of private international capital markets, and over time these markets have come to dwarf those for goods and services. The removal of capital controls from many advanced economies in the 1970s and 1980s, after the termination of the Bretton Woods fixed-exchange rate system, opened these economies even more to global markets. While labor flows never became as free from government intervention as those of capital and goods, largely because of immigration restrictions in the developed world, there was some return to the greater mobility of labor that had prevailed in the late-19th-century period of high globalization. The extension of the European Union (EU) to new sets of countries and new issues also characterized this integration movement. By the 1980s, most of the advanced industrial world, except the former Soviet Union and its satellites, had joined the world capitalist economy. Globalization was only partial in this period.

From the 1980s onward, the rest of the world's countries began opening their markets and borders even more and joining the world economy. The end of the Cold War, the collapse of the former Soviet Union, and the economic reform movement in China all added impetus to the globalization process, as the former communist countries and China all began to integrate their economies into the global capitalist one. In the 1990s and early 2000s, the world's developing countries began joining this world economy and its international institutions in large numbers. Most countries in the world now belong to the World Trade Organization (WTO), the IMF, and the World Bank. Democracy was also spreading throughout much of the world at this time. Countries also increasingly chose to join a wide variety of international organizations and treaty systems. Conventions on labor standards, human rights, environmental issues, and other areas became widely adopted by countries around the world. By 2009, most of the world's countries had become part of a global system that had become increasingly organized around a series of common principles and values governing numerous issues. Economic shocks, such as the 1998 Asian financial crisis and the 2008 financial crisis begun in the United States, spread across the globe rapidly. The second period of globalization had arrived. A major issue is to what extent globalization was purposefully constructed, rather than unintentionally realized. Whether globalization is reversible or not is also at issue.

Globalization has been studied from two general perspectives. On the one hand, the causes of globalization have attracted much attention. Why has there been a movement toward increasing contact and interaction between states and peoples all over the globe? Why has this occurred more at certain times than others? Why have some regions or peoples been more involved in this process than others? Why have certain aspects of globalization such as international trade moved faster than others? Many questions have been raised about globalization as an outcome that in itself needs explanation. On the other hand, the impact of globalization has also received great attention. What effects does this process of increasing contact and interaction among states and peoples have on the economy, politics, society, and culture? Are these effects large or small? Negative or positive? Delineating the myriad effects that globalization might have and developing evidence for these has been central to the literature. Once we understand more about globalization, we can then address the question of how to manage the process and its effects.

The Causes of Globalization

A question of much import has been how globalization has developed, and why it has developed the way it has. If one adopts the secular view of globalization, then its advance can be explained by how fast technological change occurs and spreads. In periods where such change is very rapid and when the adoption of new ideas and techniques is quick, globalization will progress rapidly. Many seem to feel that this was the case in the mid- and late 19th century, when critical innovations such as steam power, railroads, and the telegraph changed political, social, and economic life around the globe. Spread in part by military competition and imperial control, these technologies helped integrate distant regions of the world. Technological change did not stop in the interwar period when globalization was in retreat; indeed, war tends to spur such innovation. But the spread of such technologies in this period was much more limited since many borders were closed and countries guarded their new secrets, which were often tied to war fighting. The development of nuclear weapons is a case in point. The period after World War II has again been viewed as one when rapid technological change was fostered by governments, when the spread of new processes and products was unhindered, and when the adoption of innovations at least in the noncommunist countries was embraced by states. Globalization thus sped up after 1950 because technological change accelerated.

It is important to note that technological change is not just about the economy; such changes have major ramifications for society, culture, and the polity. Electricity, the telephone (not to mention the mobile phone), radio, television, refrigerated shipping containers, automobiles, airplanes, and now the Internet—to name just a few of the major innovations of the last century and a half—have affected social, political, and cultural life in enormous ways. They have all contributed to globalization by reducing the cost of transportation and communication among individuals across the globe. In this secular story of global integration, countries can try to impede innovation and slow its spread, but they cannot ultimately halt such change. To the extent that technology reduces transportation and communication costs, technological innovation provides an irrepressible force for globalization. The pace of globalization may be affected but not its forward movement.

In the cyclical view of globalization, other factors are often cited as being of great importance to the changes we see. One theory focuses attention on the salience of having one country lead the globalization process. Hegemonic stability theory (HST), developed in the 1970s and 1980s, argued that a single world leader was necessary for an open, stable world economy. Without such a leader, even technological change would not bring globalization. Two versions of this theory exist. In one, the hegemon, who has overwhelming capabilities in all areas relative to other countries, exercises a benign influence over the rest of the world by organizing and paying the costs for such openness and stability. It plays the role of leader and others benefit from its efforts. The United Kingdom (UK) and the United States are cited as two examples of this in the mid- to late 19th and mid- to late 20th centuries, respectively. The loss of British leadership capacity and the failure of the United States to exert leadership in the interwar period were seen as the leading causes of the closure and turmoil of that period. Without such benign leadership, the world would not be able to integrate and countries would remain in an isolationist stance trying to beggar their neighbors. The hegemon solves the problem of trust and competition for the smaller states and thus allows everyone to emerge better off through its provision of global public goods. It acts as lender of last resort and tries to provide an open market in difficult economic times. By creating international institutions to help countries in difficult times, the hegemon can also promote stability and bind itself to reassure other countries of its benign motives. Globalization requires leadership from the world's strongest power. Some note that perhaps a small group of powerful countries could provide this same leadership, but concerns about disagreement or buck passing within this group have shed doubt on this view.

In a second view, the role of the hegemon is not seen as selfless. Here the global leader seeks to maximize its own interests and uses its superior resources to bribe and coerce others into following it. British and American hegemony is seen as a coercive process by which the country imposes its will on others; the hegemon gains, while everyone else submits. In the British case, imperialism combined with open markets and the dominant role of the pound helped it induce others open their markets and join the international trading and financial system. Even international law as developed in this period served to enable globalization as the European powers used it to force open the colonies. British domination ran into increasing difficulty, however, as other countries grew and became competitors. Once Britain could no longer dictate the terms of engagement with the world economy, regional blocs developed around the strongest powers. Britain could no longer force countries, including the United States and Germany, to follow its preferred practices, such as joining the sterling system or opening their markets to international trade. The decline of British hegemony then ended the period of globalization as powerful rivals developed their own spheres of influence.

In the case of the United States, its dominance in all domains after World War II helped it establish a system of international governance that has maintained American hegemony for over 60 years. Unlike the British earlier, the Americans set up a series of international institutions to embed the U.S. global order. The GATT/WTO, IMF, World Bank, and regional development banks have, among others, been a means for the United States to exercise its domination without having to do so directly. These institutions have pushed countries to adopt capitalist practices, to open their markets to flows of goods, services, and capital, and to follow the so-called Washington Consensus in other economic matters. Allowing or encouraging countries to become indebted and then imposing conditions on them when they are unable to service their debts has allowed international institutions to open the developing world to the global economy. More recently, pressure for democratization has also come from these institutions. Thus, the American method for controlling the global system has differed from the British one, but it has also involved coercing others to join the global economy. The use of soft power has been perhaps more evident in the case of the United States. But American military intervention around the world has been plentiful, and this view has served to enforce its global vision. In this view of coercive hegemony, first the UK and then the United States gained much from their exalted position at the center of the global system; ironically, however, in both cases their decline seems to be connected to their leadership role. This view of the malign hegemon is associated with both a realist and a Marxist approach to international relations. Substantial debate exists as to whether the coercive or benign view of hegemony is more evident. This debate includes controversy over the extent to which the countries other than the hegemon have been hurt or helped by this globalization process.

A more constructivist account of the globalization process would focus on the way ideas and norms about the global community and modernity shape state behavior. It would identify the dominant ideas of the past 60 years and how these have been legitimated and spread throughout the system. Beliefs in liberal democracy, human rights, capitalism, and neoliberalism, for instance, have diffused across the globe; these institutions and practices have become legitimated and viewed as the "modern" way to manage a country. Countries thus aspire to introduce and develop them in order to be a part of the global community. International institutions embody and diffuse these ideas, as do norm entrepreneurs from global civil society. The focal effect of these ideas, however, is to reduce diversity and contribute to homogenization. Scholars are often divided as to whether this normative globalization is beneficial or not. Clearly, the spread of some ideas (e.g., democracy and human rights) is seen in a more positive light than the spread of others by some scholars (e.g., capitalism and neoliberalism). There is also debate over how certain ideas come to be dominant. Is it through coercion or inducement by a powerful country, such as the United States, that others espouse these ideas? Is it more the zeitgeist of the time that makes certain ideas fit the situation better? Or is it the impact of international political entrepreneurs who give legitimacy and moral suasion to certain views, independent of countries?

Another reason given for the rise of globalization is much less intentional. The uncoordinated actions of countries over time have simply resulted in an increasingly open and interconnected system. No country desired such a result, but each one acting on its own best interests adopted policies that resulted in a globalized world. This view has support since the reasons for and sequencing of many countries' decisions to globalize were nationally distinctive. The advanced industrial countries liberalized trade and then their capital markets back in the 1960s to 1980s. The EU was a major force for movement in this direction; however, European countries joined and expanded the Union for reasons other than globalization, such as domestic politics or international security. China then launched its economic reform program in the late 1970s and early 1980s for internal reasons, and political reforms have not been forthcoming. The former communist countries only moved toward a more open and capitalist economy after the end of the Cold War and the demise of the former Soviet Union in 1991, and for many of them, joining the EU had important domestic political and international security motivations. In the developing world, democratization often came before economic liberalization. Many poor countries did not liberalize economically until the late 1990s and later. Indeed, the most striking is the different paths that countries have taken since the 1970s. The manner in which and the degree to which countries are integrated into the world system are remarkably varied. While the overall trend has been toward greater contacts and openness, there has been substantial diversity in how states realized this. For instance, English is spoken much more in most countries today than 30 years ago, but the penetration of English varies greatly from country to country. In this view, this uncoordinated process has been driven by different causes in each country. Globalization has developed differently for each; and the fate of globalization lies with the (somewhat independent) decisions of countries. The reactions to the global financial and economic crisis of 2008 to 2009 underline this view since governments have failed to coordinate and have instead responded distinctly.

Consequences of Globalization

In addition to the debates over the causes of globalization, there is much debate about its consequences. There are at least three distinct views in the literature: (1) Whereas some see its consequences as negative in all aspects, (2) others see its effects as positive on the whole, and (3) still others stand firmly in the middle and see it as having both costs and benefits. Globalization is said to have many different effects. The approach here is to look at a number of those effects and sketch out the different opinions scholars hold on each. At this stage, there is little consensus on its effects. Substantial literatures exist on whether (and how) globalization has affected economic growth rates, poverty, inequality, democracy, conflict, state capacity, policy and institutional convergence, cultural diversity, volatility and the diffusion of crises, and balance of power between capital and labor. Assessing the impact of globalization implies that one must hold other factors constant, which is a difficult task. Since globalization is an ongoing process, definitive answers to these questions cannot be given. What is striking, however, is the wide range of views and evidence that exists for each of these outcomes.

Economic Growth

Globalization is alleged to have had a variety of effects on economic growth. Growing integration of national economics into a wider global one has usually been achieved through trade liberalization and the pursuit of foreign investment, and sometimes through capital market liberalization. The consequences of these policies for economic growth are still debated. Some research supports the idea that all of the policies increase growth; there is probably more evidence that trade and foreign investment support growth than does capital market liberalization. Research also shows evidence that countries that globalized in the past 30 years grew faster than those that did not. Many expected that the integration of developing countries into the world economy would promote economic convergence—that is, it would make the poorest countries grow faster and hence catch up over time with the richest. There is some evidence that this has occurred in parts of the world—mainly Asia—but not much support for it globally. The North–South divide does not seem to have closed much over the past few decades. In part, this is because the North has had faster growth rates than some developing countries. For a wide range of countries, however, globalization does not seem to have fostered faster growth. Scholars point out that for parts of Latin America and Africa growth rates were higher in the decades before the 1980s than after. The impact of trade on growth for many countries has also been ambiguous; some, especially in Asia, seem to have gained from it. But many others, in Latin America and Africa, for instance, do not. Foreign direct investment has tended to flow to the most rapidly developing economies and so it is associated with fast growth, but this may not be causal. Finally, there is much skepticism that capital market liberalization for developing countries is or has been good for economic growth. Some of the problems associated with globalization, such as inequality and volatility, seem to have affected its capacity to deliver growth.

Poverty and Inequality

A large debate centers on the impact of globalization on poverty and inequality. World Bank data suggest that from the early 1980s to roughly 2005, poverty declined globally. The number of extreme poor (those living on less than $1.25/day) decreased from 1.9 billion in 1981 to 1.4 billion in 2005, which is equivalent to a decrease from 50% of the developing countries' population to 25% thereof (*World Bank Annual Report 2009*, p. 61). Also, for example, in the East Asia and Pacific regions, the percentage living on less than $2 a day fell from 69% in 1990 to 25% in 2007. Nevertheless, 2.5 billion people in the world still live on less than $2 a day (*World Bank Annual Report 2008*, pp. 14, 34).

Other measures of the quality of life in developing countries have also improved: Literacy rates, life expectancy, and infant mortality, for example, have ameliorated in most regions except where HIV/AIDS has struck hardest. Was this decline due to globalization? Some countries that have lowered trade barriers and promoted foreign investment, such as China, Taiwan, and Vietnam, have experienced significant declines in poverty in the past decades. However, even for these countries, it is not clear that within-country inequality has been reduced; the growing rural–urban divide in many developing countries suggests rising inequality. Many other countries have seen little change in poverty levels, and inequality has either remained fairly constant or risen.

The World Bank associates greater globalization with increased growth, declining poverty, and falling inequality—especially globally. But other data call this optimistic scenario into question. A number of studies of trade have shown that its liberalization in developing countries does not reduce poverty or inequality. The so-called skill bias associated with trade and foreign investment today often means that more highly skilled workers gain more from international integration than do low-skilled ones, which tends to exacerbate the degree of within-country inequality. Interestingly, inequality seems to also have risen recently in the developed countries. This outcome was not unexpected. Standard models of trade suggest that in rich countries, high-skilled workers should gain the most from globalization and low-skilled ones should be the losers; this is the opposite of what should happen in the poor countries. This seems to have occurred in a number of developed countries, although it has probably been tempered by the redistributive effects of the welfare state. Inequality has thus risen in a number of developed countries.

The more surprising outcome has been that low-skilled workers in the developing world have not done better. While there are now numerous reasons articulated for this result, it is a problem for poverty reduction and equity in the developing world, especially since these countries by and large do not have well-developed welfare states. The impact of financial crises and their diffusion in a global system are also concerns with regard to poverty and inequality. Some argue that such crises are more likely, more intense, and spread more broadly in a globalized world. Others point out that with an open economy domestic and international shocks may balance each other out and actually make the system more stable. There is little doubt, however, that these financial shocks can have negative consequences for countries. The Asian financial crisis of the late 1990s and the recent global financial crisis have both increased poverty and inequality. How governments shape the globalization process in their countries seems to be important in influencing the way globalization affects poverty and inequality. This suggests that countries will have quite different experiences with globalization.

Politics and Democratization

Globalization has also been credited with having effects on politics. Some argue that it has helped spread democracy and put pressure on leaders to democratize. Democracy has certainly increased globally since the 1980s. But has this been due to globalization? Globalization through an open economy or through international pressure generated by international institutions or norm entrepreneurs may induce leaders to adopt more democratic forms of governance if they want to be members of the international system. Some argue the reverse: Democratization was necessary for the change in policies in many countries that led to greater globalization. The causal connection between democratization and globalization is much debated. There is also the difficult question of whether democracy and globalization are compatible. Political regimes practice democracy within their boundaries, and publics see their governments as responsible for the outcomes they experience. But in a globalized world both of these may be compromised. Democracy may have little meaning if the most important outcomes result from global forces and if governments can do little to affect these forces. The supposed democratic deficit in the EU is one manifestation of this problem. The growing number of international institutions may also be a related concern. Many of these institutions are established to facilitate international cooperation and prevent countries from pursuing beggar-thy-neighbor policies. But if they take decision-making power away from governments, they may undermine democracy at home and the public's faith in it. Others have argued that democracy may be enhanced by such international institutions. The spread of democracy and this wave of globalization seem to have gone hand in hand; how countries react over time to a highly globalized world may or may not be propitious for democracy.

Conflict

Globalization may have an impact on conflict as well. The frequency of international wars has declined steadily since the World Wars, but civil wars had risen in frequency until recently. Some research suggests that aspects of globalization can reduce conflict, especially interstate war. Research has shown that increased trade among countries, increased foreign investment, and trade agreements are all associated with less international conflict. The so-called international capitalist peace is one example of this argument. International institutions also seem to have a similar relationship: The more of them a country joins, the less likely it is to get into military conflicts. There do not seem to be similar connections for civil war. On the other side of the ledger, however, as countries trade and invest more with each other, they become more likely to have trade and investment disputes. Over time these can exacerbate, if not create, conflicts. Furthermore, as distant countries and peoples are brought into ever closer contact, the potential for misunderstandings and disputes rises as their different cultures, values, and beliefs clash; the so-called clash of civilizations is more likely in a highly interdependent world. As their interdependence increases, countries also become more vulnerable to the actions of other states. International economic ties have the potential to be used as political leverage. An interesting example is the case of China and the United States in the early 2000s. The growth of trade and investment ties between the two countries has resulted in very large and imbalanced flows of capital and trade from China to the United States; these economic linkages can serve as potential political levers for both sides. The relationship has become much more complex and fraught with political stress. Some think that these ties will dampen any propensity to engage militarily; others see them as potential sparks for a future conflict. Globalization may reduce tendencies toward conflict by making it more costly since countries will have to break their economic ties if they fight; or it may induce conflict as they have more and more linkages that can lead to disputes.

State Capacity to Govern and International Cooperation

Questions have arisen about the impact of globalization on state capacity. That is, does globalization weaken countries and erode their capacity to govern? States are often judged by their capacity to manage their economies, to respond to crises, and to provide public goods such as health care, national defense, and education. As they become increasingly intertwined in a global economy, states may lose the ability to manage their economies. The smaller the national economy relative to the world economy, the more likely is this outcome. Countries will experience the externalities of other countries' policies but will often be unable to manage these negative consequences. Losing control over monetary policy, especially if a country opens its capital account and lets its currency float, is a well-known example of this. Tax competition among states in a globalized world is another concern, since taxes provide the resources for governments to provide public goods. The race-to-the-bottom phenomenon in general that is associated with globalization can signal a country's increasing loss of control over its economy and perhaps its polity. As noted above, for democracies this loss of control can be especially worrisome as it may induce a loss of public faith in the government and in democracy overall.

On the other hand, international cooperation and international institutions may help states alleviate the problems associated with globalization. State behavior can create negative externalities for others even in a system that is not very globalized. But without globalization it may be very hard for states to cooperate to address these externalities. Globalization may awaken states to the need for organized cooperation. It may induce the creation of international institutions that help states deal with these externalities and thus provide greater political capacity than otherwise. One can think of the EU in this light. The states within the EU may now have greater capacity as a group to affect their economies than they did before joining the EU. Individually they might not be able to manage the pressures of globalization but with an institutionalized cooperative regime they may be much better able to do so. The recent interest of countries such as Iceland in joining the EU after the 2008 financial crisis suggests such an outcome. Globalization no doubt creates or exacerbates problems that do not respect national borders, but it may also contain pressures that allow countries to better respond to those problems.

Convergence Among States

Globalization has also been associated with policy and institutional convergence among states.

Some have suggested that as countries open their economies they become more likely to adopt similar policies and institutions. This convergence process can be driven by different pressures. Some attribute it to the power of the world's hegemon; some, to the increased vulnerability of states to the pressures of international institutions; some, to increased competition among states; others, to a learning process undertaken by countries. Whatever the cause, this convergence seems to have occurred in economic policy and institutions. The widespread adoption of neoliberal policies, or the so-called Washington Consensus, has been notable since the late 1980s. Other examples are the decisions by many countries to create independent central banks and to allow their currencies to float. This process has been remarked in other areas, from the creation of bureaucracies to deal with science and technology issues to the turn to democracy itself. Globalization may narrow the choices that states have and pressure them to adopt similar institutional forms and practices. To the extent that these forms and practices result in better outcomes, this process may be welcomed. However, if they are not productive for states, then this convergence process may be a negative for governments everywhere.

When this convergence pressure extends to cultural and social life, it is often seen as a negative force. Losing the cultural and social diversity that are associated with distinct regions, ethnic groups, languages, and nations is an often remarked effect of globalization. The disappearance of many languages and the increasing use of English are two such examples. For some, this is a benefit since transaction costs are greatly reduced if all people speak the same language, and the dream of a common global language has a long history. The pressure on countries or peoples to curtail certain practices and rituals that are not considered "modern" or "civilized" is said to be growing as well. Debate rages over whether countries have the right to condone all types of practices within their borders or whether there exist certain minimal standards that all countries must follow. And the end of some practices, such as slavery, has been widely applauded. The decrease in both communication and transaction costs in a globalized world has obviously helped intensify such convergence. External pressures to conform to global norms may be stronger today than ever but such pressures have always existed. Today they are less likely to be imposed through the use of force or conquest but that does not mean these pressures are less powerful.

Scholars also worry that a global system is an increasingly interconnected one where problems in one country or region can spread more rapidly, forcefully, and widely than in a less interconnected environment. The increasing ease of transportation, for example, makes the spread of disease much faster and perhaps deadlier. The communications revolution does much the same for new ideas and fads. Information cascades are more likely in such systems. This means that instability may be greater as well. Small changes to the system radiate out in all directions and can lead to large consequences. There is a concern that crises are also more likely and more contagious in such a globalized world. The Asian financial crisis of the late 1990s and the recent global financial crisis of 2008 to 2009 are two such examples. Some have argued that the world's largest economies have experienced a "great moderation" since the 1970s with a dampening of business cycles and decreasing inflation due partially to greater international exposure. Others see an international economy beset by rising volatility over time, as crises multiply and spread. Globalization may stabilize a system by spreading problems out and bringing counterbalancing forces to bear, but it may also destabilize a system by creating a network of linkages among all countries and across all sectors that transmits and magnifies problems.

Domestic Economies

In addition to the question about which countries gain and lose from globalization, there is the issue of which groups gain and lose domestically. Many see globalization as strengthening the hand of capital owners and right-wing political parties as opposed to labor owners and left-wing parties. With a global economy where capital can move relatively freely, it becomes advantaged domestically. Globalization changes the domestic balance of power between business and labor. When owners of capital do not like a government policy or fear labor has become too strong, they can often move to a new political environment that is more accommodating. This process can then generate race-to-the-bottom pressures as governments try to retain capital in their country by deregulating and adopting more business-friendly policies. Some claim that globalization has helped undermine labor unions and other forms of labor market organization. This process can then influence the political system. If right-wing parties represent business interests more, they may be strengthened by globalization. Left-wing parties and their programs may seem increasingly unattractive as they "scare" capital away and perhaps slow down growth. Evidence for this effect is mixed. Some claim that left-wing parties and governments actually do a better job wooing capital and thus are not disadvantaged by globalization; they may even make better partners for international business than right-wing parties. Some scholars see no effect of globalization on domestic politics; they claim that domestic political institutions filter and shape such external pressures so that their ultimate effects are a function of domestic politics. Preexisting domestic institutions and practices then govern whether and how globalization affects internal politics. The impact of globalization on political power within and between countries is a topic of debate and great importance for future research.

Limits of Current Research

The consequences of globalization could thus be widespread. They might affect economics, politics, society, and culture. Existing research points out that globalization has affected economic growth rates, poverty, inequality, democracy, conflict, state capacity, policy and institutional convergence, cultural diversity, volatility and the diffusion of crises, and balance of power between capital and labor. But it provides little consensus on the direction or magnitude of these effects.

Globalization and the Future

Given the many consequences that globalization might have, one question that is often raised is whether countries can manage its impact. This is especially important to the extent that globalization has negative consequences. Many types of problems that have been endemic to countries for centuries have taken on a new perspective in a globalized world. Problems that were national have become transnational. The ability of countries to realize many of their basic goals has become more and more tied to the actions of other countries. Economic prosperity domestically often depends now on international trade and capital movements around the globe; national security from all sorts of threats depends on the behavior of other countries and nonstate actors; environmental conditions rely on the behavior of other agents globally; and public health relies more and more on transnational factors and the behavior of international agents. Countries have always faced such problems; it is just that their impact and resolution depend more and more on agents outside the state itself. Globalization is intimately connected to this process. By increasing contacts and interactions among countries and peoples, it helps make many issues transnational in character, rather than national.

Given that governance is largely the domain of nation-states, is there anything that can be done to address these transnational problems? One country can rarely dictate another's economic, environmental, security, or public health policy, for example. But to deal with transnational problems, coordination of policies may be the best way forward. In particular, cooperation where countries coordinate their policies to arrive at mutually preferable outcomes may be essential. Globalization may facilitate cooperation; it may make the costs of failing to cooperate so high that countries are more willing to try. But it may also make the stakes of cooperation much higher and thus render it less likely.

International institutions may be one way to address these concerns. In a globalized era, some see such institutions as essential for dealing with transnational problems. Countries alone cannot successfully deal with them; they require changes in the behavior of other states. International institutions that states voluntarily join and comply with may help them. Such institutions may help states realize cooperative outcomes; they may provide transparency and lower transaction costs for negotiating solutions to transnational issues; they may embody global norms and practices that allow states to identify focal points for cooperation and/or help them enforce compliance with international norms and cooperative agreements. These are the potential benefits of global institutions. They may have costs as well. Not all international institutions function adequately; some have serious internal defects that cause them to operate poorly. Others are deadlocked by internal divisions. Some are dominated by one or two states that coerce others to adopt their preferred norms and practices even if these are not very beneficial for the others. International institutions cannot be seen as a costless and efficient solution to all transnational problems. But many of them have shown some ability either to prevent transnational problems for worsening (e.g., the WHO and the WTO in times of crisis) or to allow states to cooperate in order to better manage problems in a globalized world (e.g., the EU).

A related issue is concern over globalization's future. If one adopts the secular view of it, then little chance exists that the increasing contact and interactions among peoples and countries can be reversed. Indeed if there is technological progress, there is likely to be increasing globalization; more and more contact and interaction are inevitable as technology brings people in the world closer together. On the other hand, the more cyclical view suggests that globalization comes and goes in waves depending on political, economic, and social reactions. We may now be experiencing the height of this period of globalization. Forces for a backlash against the pressure exerted by increasing contacts and interactions among peoples and states may be gaining strength. The economic and financial crisis of 2008 to 2009 may be a prelude to this type of backlash. In this view, governments can respond in ways that curtail or reverse globalization; they can protect their economies, close their borders, end participation in international institutions, and take other steps that seal them off from the rest of the world. The period between 1914 and 1945 is one example of this. The question, of course, is what would be the costs and benefits of such a course of action and would governments be able to achieve any kind of meaningful autonomy or autarchy, given the technologies we now have and the public expectations about their connections to the rest of the world. Whether globalization is reversible or not is a question of great import; the costs of doing so would be a key factor. More realistic, perhaps, are questions about whether globalization can be slowed down or managed in ways that are better for all concerned.

In conclusion, debate rages about globalization. A number of theories exist about its sources. Technological change, international political hegemony, global normative convergence, and/or independent country reactions have all been cited as major causes of globalization. Globalization's effects have been debated in even greater intensity and breadth. Substantial literatures exist on whether globalization has affected economic growth rates, poverty, inequality, democracy, conflict, state capacity, policy and institutional convergence, cultural diversity, volatility and the diffusion of crises, and balance of power between capital and labor. Agreement on the nature of its effects on any of these is scarce. For some, the balance of its impact has been negative in almost all domains; for others, it is seen as largely positive, especially in the economic and political areas; and for yet others, it has costs and benefits that are hard to calculate and summarize overall. One fact not in doubt is that globalization is an important feature of our world these days; it seems to have consequences for all domains of economic, political, social, and cultural affairs. How to manage it in order to make states, peoples, or the world better off is thus of central importance.

Helen V. Milner
Princeton University
Princeton, New Jersey, United States

See also Capitalism; Colonialism; Conditionality; Democracy, Theories of; Development, Political; Governance; Hegemony; Human Rights, Comparative Perspectives; Human Rights in International Relations; Inequality, Economic; International Monetary Fund (IMF); International Organizations; International Trade; Multinational Corporations (MNCs); Neoliberalism; Power; Realism in International Relations; Trade Liberalization; World Bank; World Trade Organization (WTO)

Further Readings

Bhagwati, J. (2004). *In defense of globalization*. New York: Oxford University Press.

Cox, R. W. (1987). *Production, power, and world order: Social forces and the making of history*. New York: Columbia University Press.

Eichengreen, B. (2008). *Globalizing capital*. Princeton, NJ: Princeton University Press.
Frieden, J. (2008). *Global capitalism*. New York: W. W. Norton.
Garrett, G. (1998). *Partisan politics in the global economy*. New York: Cambridge University Press.
Greider, W. (1997). *One world, ready or not: The manic logic of global capitalism*. New York: Simon & Schuster.
Hirschman, A. O. (1980). *National power and the structure of foreign trade*. Berkeley: University of California Press. (Original work published 1945)
Hoogvelt, A. M. M. (1997). *Globalisation and the postcolonial world: The new political economy of development*. Houndmills, UK: Macmillan.
Keohane, R. O., & Milner, H. V. (Eds.). (1996). *Internationalization and domestic politics*. New York: Cambridge University Press.
Kindleberger, C. P. (1973). *The world in depression*. Berkeley: University of California Press.
Milanovic, B. (2007). *Worlds apart: Measuring international and global inequality*. Princeton, NJ: Princeton University Press.
Polanyi, K. (1957). *The great transformation: The political and economic origins of our time*. Boston: Beacon Press. (Original work published 1944)
Robinson, W. I. (2004). *A theory of global capitalism: Production, class, and state in a transnational world*. Baltimore: Johns Hopkins University Press.
Rodrik, D. (1997). *Has globalization gone too far?* Washington, DC: Institute for International Economics.
Stiglitz, J. E. (2002). *Globalization and its discontents*. New York: W. W. Norton.
Strange, S. (1988). *States and markets*. London: Frances Pinter.
Wolf, M. (2004). *Why globalization works*. New Haven, CT: Yale University Press.
Woods, N. (2006). *The globalizers*. Ithaca, NY: Cornell University Press.
World Bank. (2008). *The World Bank annual report 2008: Year in review*. Washington, DC: Author.
World Bank. (2009). *The World Bank annual report 2009: Year in review*. Washington, DC: Author.

GOVERNANCE

Amid the debate on globalization and its influence on domestic politics and institutional arrangements, it is intriguing to note the resilience of national institutions, policy styles, and perceptions of the role of the state in society. True, many countries that displayed significantly different governance arrangements a few decades ago today present a more common governance model. Even so, however, the legacy of—or more correctly, perhaps, the continued support for—traditions of statehood, governance, and collective action loom large. These traditions can be seen as value systems that have evolved over an extensive period of time, continuously reproduced in public policy and public discourse. If there is strong resilience in institutions and policy, this in part can be attributed to strong political and administrative tradition. This entry examines the nature of the contemporary state and then examines the different governance models as means of solidifying policy preferences.

Although it makes sense to think of these value systems as traditions, they can mostly be traced back to specific political contexts and policy choice (see, e.g., Francis Castles, 1993). For instance, the Asian "developmental state" (Chalmers Johnson, 1982) and the Scandinavian and continental European welfare state models (Gøsta Esping-Andersen, 1990) evolved through a series of political choices about the role of the state, the market, and civil society. Against this backdrop, the purpose of this entry is to explore the linkages between policy choice and models of governance. Conventional wisdom holds that policy choice is largely a product of institutional arrangements. The immediate interest in this entry is to see to what extent those institutions, in turn, can be related to a more overarching set of values and beliefs about the role of the state. (*Value* refers to everything that is considered to have a value, be it material, ideological, ethical, or other, by most people/citizens.) Should that be the case, we will be well under way to connect a typology of governance arrangements to broad categorizations of policy. There has been some work done on the relationship between different types of policy, drawing on the familiar typology developed by Ted Lowi, and the institutional arrangements of the state but this is a vast research field and much more work remains before we have a more complete account of the linkages between governance arrangements and policy styles.

This entry then looks at a couple of different governance models as a form of institutionalization of overarching policy choices in terms of constitutional and institutional arrangements. The underlying premise of this entry is that the institutions framing policy choice are themselves "locked in" by fundamental policy choices concerning the role of the state in society and its centrality in governance. The predominant governance model of a particular country is not a given but reflects entrenched notions about the role of the state, the market, and civil society.

These values and perceptions become deeply institutionalized over time and elevated into political culture above and beyond day-to-day policy making. Even so, however, when observed over a longer stretch of time distinct changes in these value systems can be observed. The role of government in the United States changed fundamentally from Lyndon Johnson's Great Society project in the 1960s to Ronald Reagan's neoliberal policy style in the 1980s; in Britain, Labour under Tony Blair and Gordon Brown looked significantly different from Labour under Harold Wilson; in Japan, the relationship between the Liberal Democratic Party (LDP), the bureaucracy, and the corporate sector has changed profoundly since the 1960s; and in Scandinavia, the welfare state today has very different organization and modus operandi from those of the 1960s or 1970s, and so on. Thus, however entrenched and institutionalized these values may be, they do nonetheless change, sometimes abruptly, sometimes more subtly.

The hypothesis that specific cases of policy making are embedded in more overarching values and norms, perhaps even a social theory, concerning what should be the role of the state in society opens up for potential theoretical circularities. It could be that the more general values are simply defined as the sum of specific policies, in which case the argument about embeddedness and institutionalization becomes circular. And, even if different specific types of policies can be theoretically linked to different governance arrangements, a causal mechanism is still needed to explain how and why that would be the case. A more plausible way of looking at such relationships would be to suggest that policy and governance tend to foster and reinforce each other. For instance, "big government" service models tend to emphasize redistributive policies and equality and equal treatment in service delivery, while neoliberal policies emphasize economic growth and a customer-oriented, market-like model of public service. Thus, different aspects of public policy and public administration practices are closely related.

This entry first looks at four broad dimensions of governance that relate to different aspects of the role of the state in society. It then discusses three different models of governance: state-centric, liberal-democratic, and network-based governance. The third section merges these two analyses to elaborate the features and logics of the three governance models. Since issues related to institutions, institutionalization, and deinstitutionalization are central to the present analysis, these issues are discussed before a concluding section sums up the entry.

Dimensions of Governance

Nation-states around the world offer a wide variety in terms of how they organize their governance. Perhaps the most obvious and striking differences relate to the fundamental aspects of state organization in terms of their institutional arrangement. Thus, there are federal states and unitary states, centralized states and decentralized states, and states with a long tradition in regulating markets and social behavior while other states see their role as enabling and facilitating markets and free individuals.

These institutional arrangements are highly path dependent. They evolved as constitutional systems defining the model of governance, which were believed to reflect popular expectations and the requirements of a political system to govern the country. Thus, for instance, American or Canadian federalism was the model that was believed to provide the optimal balance between national interests and regional autonomy in those vast territories, while the unitary state was the preferred arrangement in smaller European countries. These are complex issues to capture analytically because they are shaped by the history of the country and its unification process, its state tradition, the homogeneity of the population, and its geopolitical context.

One strategy to approach these issues is to relate them to a set of dimensions of governance. If we conceive of governance as the pursuit of collective interests, with the state as a coordinating and enabling actor, we can clearly see that governance is both a process that is endogenous to the state—division of labor, competencies, jurisdictions, and resources—and a process in which the state interacts with other actors in society. Therefore, we will want the dimensions of governance to capture both the internal and external aspects of governance. Below, four such dimensions of governance are briefly outlined. These dimensions describe different relationships between the state, the society, and the market.

The governance dimensions speak to very basic issues about the organization of the state and its role in society. Therefore, it should not be surprising to find that they are normatively charged and are frequently topics of political disunity in many countries.

Popular Representation Versus Institutional Autonomy

The first governance dimension measures the degree to which the state is responsive to popular demands and expectations, and it places that responsiveness in juxtaposition to a high degree of autonomy for the state's political and administrative institutions. For governance to qualify as democratic, it is essential that the state responds to popular demands and incumbent political parties are allowed to shape the policy agenda and to engage in public discourse. At the same time, however, institutions need to have some integrity and autonomy, both in relationship to partisan ideologies and to direct popular pressures in order to be able to fulfill their role in the policy process. Institutions that develop close and bilateral relationships with parochial societal interests can develop networks as alternative instruments of policy implementation, but they will have problems in implementing policies that discriminate among policy targets. This is, thus, a dimension where both extremes are unsatisfactory from a governance perspective; democratic governance is contingent on both popular input on public policy and some degree of institutional autonomy allowing those institutions to pursue policies that might not enjoy popular support but are necessary to the state.

Another example of the complexities associated with this dimension can be seen in the relationship between politicians and bureaucrats. On the one hand, we expect the civil service to be responsive to the government of the day, implementing its political wishes without prejudice. On the other hand, we expect the public administration to be guided by values such as impartiality and by due process and public law, and a "politicization" of the administrative system is often seen as a controversial development that could compromise the integrity of the bureaucracy. Again, both extremes on the dimension are problematic in a governance perspective whereas some form of combination of the two positions fosters democratic governance or is indeed integral to governance.

Societal Integrity Versus Institutional Capacity

A second dimension of governance refers to the tension between fundamental values related to individual freedom, on the one hand, and the need for institutions to encroach on that freedom in order to ensure collective interests and values, on the other. This dilemma and the question it raises on forms of justification for political control over individuals is a major field of research and discourse in political philosophy, and the finer details of that debate are not probed too deeply here. There are obviously values related to the individual qua an individual and other values derived from the individual as a member of a polity. The tension between these two types of values plays out almost daily in a variety of policy sectors such as environmental policy and climate change, crime prevention, social welfare, and tax policy. The government's need to collect information about individual behavior has become a major issue, particularly in matters related to national security where the need for surveillance stands against personal integrity. A similar debate has surrounded means-tested welfare state programs where government agencies require beneficiaries to submit detailed information about their standard of living and income.

The degree to which government should have access to private information or regulate personal behavior is ultimately a normative issue where different interests are weighed against each other; it is not possible to objectively delineate the pursuit of individual interests aggregated into government action, on the one hand, and the rights of individuals to act freely, on the other. The point here is instead that some encroachment on individual freedom is necessary for government to govern society, while the precise extent to which such encroachment can be justified remains a normative issue that reemerges fairly regularly on the political agenda. Meanwhile, constitutional reform and institutional design have to wrestle with this issue in a more mundane and pragmatic fashion.

Institutional capacity is a key variable in a broader governance perspective. According to one perspective, contemporary society has become too complex for any single actor to provide effective steering or coordination. Even if government previously had the capacity to fulfill those roles, the growing complexity has undermined the governing capacity of the state. As a result, governance is conducted through networks, partnerships, and other interactive instruments of coordination and collective action. In terms of institutional capacity, this perspective argues that in contemporary governance, such capacity is not so much a matter of constitutional mandates as a continuous process of reproducing institutional leverage through cooperation with societal actors.

National Interests Versus Subnational Autonomy

All systems of government integrate political and administrative action on different institutional levels. Exactly how that integration is organized is, however, a matter of substantive cross-national variation. Thus, subnational political systems are granted some degree of autonomy in relationship to central government (for an overview, see Alan Norton, 1994).

The basic logic of organizing government in this way is one of efficiency and expediency, but there are also important aspects of these arrangements that speak to the development of democratic governance. The efficiency argument is fairly obvious; it simply does not make any sense to have the national government organize water and sewage system maintenance or garbage disposal service in every town and village throughout the country. The democracy argument holds that participation in local public affairs socializes citizens into the democracy form of governance. For that socialization to work, local government must have some autonomy, some jurisdiction, within which the local demos can make decisions pertaining to public matters in their locale. Also, the democratic theory of local government holds that the role of elected officials should be to cater to the interests of the populace in a nonpartisan way. Ideology and division should not be imported from national politics and brought onto the local political scene.

This somewhat cozy account of local government is troubled by the existence of national interests, which play out in the local territory. In issues such as infrastructural development, the location of windmill parks, or the deposit of nuclear waste, national interests frequently encounter fierce opposition from local governments. The question then becomes under what conditions should the central government be allowed to override the will of a local demos? The problem is larger than simple not-in-my-backyard (NIMBY) politics; it relates ultimately to the problem of two competing *demoi*: the national and the local.

Again, the two extremes of this dimension—complete dominance of national interests and extensive, unconditional local autonomy—raise serious questions in terms of the quality and, indeed, the prospects, of governance. Instead, some middle ground position characterized by dialogue between central and local government with respect for both the national and local *demoi* should provide for better governance.

As is the case with the previous dimension, this dimension, too, speaks directly to governance issues in several different ways. If central, regional, and local governments have overlapping competencies and jurisdictions, that is, if policy making involves several levels of government, the policy process will be characterized by a significant vertical dimension of resource allocation, control, and bargaining. If, on the other hand, policy sectors are largely divided among institutional levels so that central government is in control of some policy sectors, regional government of other sectors, and local governments of yet other sectors, the process becomes less complex and will not involve much interaction among institutional levels. It is thus essential to acknowledge the political nature of state–local relationships; while a formal institutional analysis depicts those relationships mainly as one of division of labor, a closer inspection of central–local interactions will uncover struggles over control, financial resources, and policy.

Another example of the governance ramifications of central–local relationships is that they define to a large extent the relationship between a city and key players in the local environment. Urbanists would refer to this as the interplay between horizontal and vertical dimensions of autonomy where the vertical dimension defines much of the preconditions for horizontal, urban governance.

Autonomous Markets Versus Regulation

The degree to which government regulates markets is often seen as a fundamental aspect of the left–right division in political life. Parties on the political right typically insist that markets work best if they are left alone; market regulation chokes the market, rendering it an inefficient instrument of resource allocation. Government regulation in this perspective distorts market mechanisms, perverts pricing, and removes incentives for risk taking and entrepreneurialism. Leftist parties, on the other hand, maintain that markets need to be brought under some public control partly in order to ensure that they do not create politically undesirable outcomes such as vast differences in wealth and life chances. Many of the areas of society that are essential to the type of social reform that parties on the political left pursue (e.g., housing, education, health care, and work), operate under market control. Therefore, market regulation is an important instrument to insert political objectives into markets.

Market regulation, as mentioned earlier, is politically controversial. While there may be agreement among political actors on long-term policy goals such as economic growth, issues become charged as soon as the role of the market toward those political goals is to be defined. Parties right of center will argue that markets left alone will deliver the highest economic growth, whereas parties on the political Left will insist that markets cannot deliver a socially acceptable economic development or that markets are simply unable to solve collective problems.

That having been said, however, there is also a fair amount of pragmatism surrounding the issue of market regulation. The American political culture has historically, albeit with few but important exceptions, heralded the notion of free markets and small government. Meanwhile, the economic history of the United States is replete with examples of publicly driven large-scale infrastructure development projects and public service delivery, according to Jeff Madrick (2009). More broadly, there is an understanding from Adam Smith's classic book in the field, *On the Wealth of Nations*, onward that markets need a regulatory framework to function properly. Furthermore, the evidence supporting the argument that nonregulated markets perform better than markets under some degree of government regulation is agnostic. The key problems seem to revolve around issues such as designing the governance of a market economy that generates growth at the same time as it solves problems of coordination and provides consumer protection and sustainable development.

This leads us to consider the critical distinction between regulation *of* and *within* markets. The regulation of markets means providing a regulatory framework defining property, contracts, and so on and a mechanism for settling disputes among market actors. Regulation in a market refers to political and administrative control of market mechanisms such as price or supply by regulating entrance to the market. The former type of regulation is nondiscriminatory, while the latter form is discriminatory and targeted to ensure collective interests in the markets such as consumer protection and price caps on specific goods provided by the market. Such public presence in the market could also manifest itself in taxes or incentives to sanction or encourage different forms of market behavior.

From a governance perspective, this dimension provides important information about the political economy of governing. Peter Hall and David Soskice's (2001) distinction between "liberal" and "coordinated" market economies suggests that the degree to which the capitalist economy is embedded in domestic governance varies considerably among different countries. In coordinated market economies, with the continental European, Scandinavian, and several Asian countries as leading examples, interactions between the state and private capital are institutionalized, making concerted governance less complicated and politically sensitive than in the Anglo-American liberal market economies. Also, market regulation is a policy instrument defining a political objective in the market. That objective could, however, also be implemented in less obtrusive ways, thereby facilitating a higher degree of concerted action among public and private interests.

As is the case with the other governance dimensions, both end points of the dimension—completely free or completely regulated markets—are neither efficient from a market point of view nor perhaps desirable from a democratic governance perspective.

Governance Models

Let us now turn to a brief presentation of a few stylized models of governance. In 2005, Jon Pierre, together with Guy Peters, outlined a theory of governance drawing on a typology of governance models. The basic assumption sustaining that analysis was that all governance arrangements have to fulfill a number of roles or functions, but different governance models accord the state and societal actors widely different significance for those roles. The authors found that both the most state-centric governance model and, at the other end of the spectrum, network-based governance had problems with defining a state–society interaction that blended institutional autonomy with openness to the interests and expertise found among societal actors. Instead, governance models between the two extreme positions on the continuum, such as the corporatist governance model, fared relatively well; they were capable of blending some indigenous capacity to steer with a fair openness in state–society exchanges.

The defining feature of governance theory, as opposed to more traditional government-centric accounts of governing, is the centrality of exchanges between the state and actors in its external environment. The three models outlined here describe different types of such relationships and interactions. However, the present space only allows for very brief accounts of complex and multidimensional governance arrangements.

State-Centric Governance

The first governance model portrays the state exercising firm control over its territory and population. The state has few contingencies to other societal actors and does not have to be concerned about any lack of institutional capacity to impose its will on society. The typical governing style of government in state-centric governance is to use policy instruments sufficiently distinct to ensure the desired change in social behavior. That includes securing tax revenues of sufficient level to sustain those projects that the state wishes to implement. The state, in this model of governance, sees little reason to engage in collaborative forms of governance or service delivery.

If we accept Joel Migdal's (2001) argument that state and society mutually constitute each other, democratic forms of state-centric governance occur where there is a societal expectation for such a leading role of the state. That means that although government will not refrain to use obtrusive policy instruments, the societal submission to the state suggests that the state does rarely have to throw much weight to impose its will on society. Whether it is because the state sees it as its role to provide distinct leadership or because society expects it, state-centric governance is built on the centrality of the state. That centrality is obviously a significant asset in terms of being able to steer and control societal development. However, it tends to entail a fairly unidirectional process of state–society exchange whereby the state executes its preferences on society. States in this model of governance easily become self-referential and arrogant vis-à-vis societal actors and may thus be unaware or oblivious of societal change.

France is historically seen as a prime example of state-centric governance although that state predominance is today somewhat less articulated. We also find examples of state-centric governance in many Latin American countries such as Argentina (not least under Juan Perón) and Uruguay. Note that state-centric governance does not correlate with the existence of leftist political regimes. The legitimacy of state encroachment on society stems from a symbiotic notion of the state embodying the people and catering to its interests like parents look after their children. In this model of governance, nationalism plays a significant role and it is legitimate, and expected, of the government to intervene in markets if it can be justified with reference to interests of the state.

Liberal-Democratic Governance

The previous model of governance is a relatively rare phenomenon in time and space. Liberal-democratic governance is a much more common creature. The Anglo-American democracies and the continental European countries belong to this category as do some of the Asian countries, particularly with regard to their post-2000 model of governance. Thus, although the fundamental features of the governance model cover a fairly narrow section of the spectrum of different governance models, we should expect to find a fair degree of empirical heterogeneity as we inspect this model in more detail.

In terms of democratic theory, liberal-democratic governance rests on a notion of a distinct separation between state and society. The state is not expected to interfere in markets and to honor values of individual freedom and personal integrity. However, the state is expected to play a key role in the regulation of markets by upholding a regulatory framework for markets and ensuring free competition. Thus, in the spirit of Adam Smith, the "regulatory state" model sees the government enforcing antitrust legislation to protect markets from cartels and other forms of improper market behavior. Thus, there is an emphasis of regulation *of* the market and an equally strong normative belief that government should regulate *within* markets. In the parlance of economic theory, it is the role of the state in this governance model to prevent market failure. It should not, however, pursue political objectives through regulation within markets.

The state in liberal-democratic governance is more deeply embedded in society than state-centric governance, yet it is able to maintain some degree of institutional integrity in relationship to policy targets and other social interests. The state is not alien to collaborative forms of governance but insists on playing the role of a primus inter pares in those arrangements. That means that government will consult with social partners in policy making but will reserve the right to make the final decisions on policy design. Thus, in some ways liberal-democratic governance could be just as intrusive on society as state-centric governance, but the mechanisms and processes sustaining that capacity of the state are less derived from institutional strength and more a result of social bargaining.

The different forms of corporatism that have been a trademark of state–society exchanges in several European countries and also in Japan fit nicely into this model of governance. This applies in particular to the tripartite arrangements between the state, private capital, and labor where the state can be both an arbitrator as well as a negotiating party. The significance of these corporatist arrangements has declined in several countries remains strong in, for instance, Denmark.

Network Governance

Some time ago now, Rod Rhodes gave new momentum to the emerging debate on contemporary governance by arguing that policy networks were the key instrument of governance, hence governance of policy sectors saw government as one of a large number of actors. Indeed, cohesive policy networks had the capacity to resist government policy and were basically capable of governing the sector themselves. Coupled with the growing literature on the governance problems that stemmed from the increasing social complexity and multilevel institutional arrangements of the political and administrative sphere, networks and other interactive forms of governance became the key focus of a large group of governance scholars.

Network governance scholars deliver a powerful argument about the lack of institutional capacity and integrity of the state. Short of those features, segments of the state embark in a range of different interactive arenas where collective solutions are designed and implemented in concert with social partners. Policy becomes reflective of the contextualized process from which it evolves; interactive governance forms have difficulties with imposing unpopular policies on those social actors and collective action becomes to a large extent defined by the least common denominator among the network participants. That said, network governance provides public officials with a multitude of social contacts that ensures up-to-date information about social change. Network governance therefore provides quicker and more effective linkages between changing societal problems and a political response to those changes.

It is difficult to think of any national context where networks do not exist. The intriguing question is how such forms of governance are integrated with more traditional, hierarchical governance and administrative arrangements and what the consequences are of blending different governance models. Network governance is predicated on substantive autonomy and empowerment for lower-level civil servants so that they can engage in meaningful dialogue with social partners. Thus, network governance challenges traditional top-down channels of communication in public sector organizations, both in terms of policy implementation and in bringing information from networks to the attention of the senior levels of the organization.

Governance as Institutionalization of Policy Choices

Following these discussions about dimensions and models of governance, this entry now presents a more elaborated account of different stylized models of governance with regard to how they score on the overarching dimensions of governance.

It is particularly challenging to look at the dynamic nature of these interactions. The dimensions of governance described earlier could easily produce a functionalist analysis of the roles of government in governance with all that entails in terms of static modeling. However, thinking about how government resolves these issues and tasks in a governance perspective yields a highly dynamic account of the role of the state since the core assumption in governance theory is that collective objectives are pursued in a variety of different concerted and collaborative process between state and society. The key assumption being explored in this analysis is that the degree to which political actors and institutions shape the governance process determines much of how specific governance arrangements relate to the three dimensions of governance.

State-Centric Governance

With the leading role of the state in this model of governance, we would expect that its institutions enjoy a high degree of autonomy in relationship to societal actors and interests. Since those institutions have very little collaborative interaction with societal actors, there is little incentive for them to make concessions on policy. Also, given the centrality of the state and its insistence on "going it alone" in terms of policy implementation and service delivery, government institutions will have the administrative and financial resources required to play those roles.

State-centric governance is also likely to display a priority for national interests over local autonomy. It is interesting to note that many of the empirical contexts that we associate with state-centric governance, again with France as a leading example, have weak and fragmented local government systems. This is obviously not a feature only of countries with state-centric governance—consider for the sake of comparison the equally fragmented local government system in the United States—but it makes sense to stipulate that state-centric societies do not develop strong subnational government as that could create or exacerbate central–local conflict.

State-centric governance easily translates into strong regulation of markets. This is because these states have a path-dependent political and administrative behavior according to which they rarely yield to the interests of societal actors. The state defines its interests and objectives in the market and implements whatever regulation is required toward those ends.

All these features of state-centric governance—strong, insulated institutions, central government control, tight market regulation, and so on—are in fact reflections of society as much as of the state. It is very difficult to think of a democratic state ascending to such societal predominance without the consent and indeed expectation from society. While this is a governance model that may appear to be highly obtrusive and insensitive to societal actors, a closer inspection of specific governance processes indicates that coercive policy instruments are not always necessary for the state to control society; societal actors anticipate the preferences of government and act accordingly. A case in point is the Japanese developmental state, which throughout much of the postwar period was a case of state-centric governance. Governance was coordinated within the tripod of the Liberal-Democratic Party, the bureaucracy, and leading corporate players. The chief ministry driving the economic modernization programs could ensure compliance for its plans by formulating "visions"; it rarely had to employ its institutional capacity to ensure corporate compliance with its policy, according to Johnson (1982).

However, as we look at the dimensions of governance, we can see that the case of Japan offers some intriguing patterns. For instance, although national interests were paramount through the history of Japan, not least during the postwar economic recovery, and although central–local relationships have been predominantly hierarchical, the relationship between central and subnational government has been replete with ideological conflict. Central government actively promoted the ascension of Tokyo to a global city status, triggering rivalries with several other metropolitan areas. National interests have translated into Tokyo interests, arguably at the expense of other cities.

The paradox of state-centric governance is thus that although the state has all the instruments and resources necessary to impose a strict regime on society, it rarely needs to use all that muscle as societal actors readily submit to more subtle forms of governing. Society expects the state to play a leading role in governance. When we bring state–society exchanges that are integral to governance theory into the analysis, it becomes clear that the more formal power bases that a state controls, the less it will have to use those forces.

The governance approach to these issues also suggests that states remain strong but that the sources of that strength have changed. State-centric governance, embedded in norms that legitimized such governance and prescribed long-term policy goals, defined overarching policy choices that continued to shape public policy for decades. Economic development has seen a distinct role for the state either as a coordinator of private business or as a leading player in research and development or as safeguard against overseas competition or all of the above. State-centrism has also had a major influence on the policy instruments typically employed in policy implementation. Again, this rather self-referential modus operandi proved to be a significant problem in times of rapid external change such as when globalization caught momentum, opening up society to new political, financial, and cultural influences.

Liberal-Democratic Governance

If state-centric governance presents a pattern of a state that enjoys extensive capabilities to impose its will on society but rarely has to use that leverage, an almost opposite pattern is found in liberal-democratic governance. In this governance model, the state has some political and institutional capabilities but the exercise of those power bases is more contested than is the case in state-centric governance. Thus, this governance model displays a medium level of institutional autonomy and an institutional capacity that is more contextualized than in the previous model.

The portrayal of the state in liberal-democratic theory is an idealized account of the state as distinctly separate from society. It could be seen as an extension of the Wilsonian model of public administration as firmly separated from the political sphere of the state. That model has lost much of its empirical accuracy in contemporary public administration and a similar development can be seen with regard to the separation of state and society. Migdal (2001) describes this development of the state nicely:

> Various parts or fragments of the state have allied with one another, as well as with groups outside, to further their goals. Those practices and alliances have acted to promote a variety of sets of rules, often quite distinct from those set out in the state's own official laws and regulations. These alliances, coalitions, or networks have neutralized the sharp territorial and social boundary that the first portrayal of the state has acted to establish, as well as the demarcation between the state as a pre-eminent rule maker and society as the recipient of those rules. (p. 20)

With the border between state and society thus compromised by parochial interests inside and outside the state, governance theory offers a new and promising approach to understanding the process of governing. Indeed, Migdal's account of the contemporary liberal-democratic state is an apt illustration of the contemporary liberal-democratic state, involved in collaborative governance and a rule taker almost as much as a rule maker. It might be argued that the state–society division that this model of governance departs from has become a normative myth. This myth may at some previous time have defined what constitutes proper exchanges between the state and its external environment, but given the complexities of contemporary governance the state–society separation has become impossible to sustain. That having been said, normative models still matter. As the rules of the political game, normative models provide definitions of roles for political actors and institutions. Those norms are also the foundation on which constitutions are written and amended.

In the liberal-democratic governance model, the state itself has moderate institutional capacity but makes up for that through collaboration with social partners. This can take place through corporatist models of interest representation and mediation but it can also be generated through more ad hoc models of joint ventures such as public–private partnerships. Also, states in liberal-democratic governance tend to define their chief role as a "regulatory state," providing market stability without necessarily being committed to extensive service delivery.

It is interesting to note that liberal-democratic governance would encompass both liberal and coordinated market economies. Government plays a distinctly different role in those two models of the capitalist economy, with a higher profile and more accentuated political presence in the market in the case of the coordinated market economy. According to Hall and Soskice (2001), there is also more legitimacy for articulating collective interests in economic development in the coordinated market economy than in the liberal market economy where government is expected to provide a regulatory framework for the economy but not to itself be a player in the market. Corporatism is clearly related to the coordinated market economy. This stronger legitimacy for market intervention in the coordinated market economy links it strongly to welfare state models and to overarching policy choices that give a higher priority to distributive or redistributive types of policy than to promoting economic growth.

Network Governance

This model of governance could be summarized as characterized by low institutional autonomy; its institutional capacity is to a large extent contingent on societal consent and a low degree of regulation or even self-regulation. In a model of governance where concerted action is the typical strategy to govern, the issue of institutional autonomy becomes moot as institutions do not require autonomy to engage social partners. Indeed, such autonomy could be an obstacle to creative problem solving since it suggests that there are rules defining the forms and conditions for institutions' exchange with their external environment.

The institutional capacity in network governance is limited; it is part of the logic of network governance that action is based on consent and therefore institutions can only implement those programs that other actors in the network agree to, which frequently comprises important policy targets. That having been said, however, when there is consensus for the execution of a policy, implementation is likely to be swift and efficient since interaction between institutions and their social partners continues throughout the implementation process.

The regulatory capacity of networks is high within the network itself but less so outside the network. In network governance theory, primary attention is on networks as instruments of self-regulation within a policy sector. In that perspective, networks offer regulation that, by definition, is reflective of the interests of the participants in the network.

This is not a model of governance that defines large-scale political projects such as welfare states or developmental states. It lacks the social inclusion and capacity to resolve conflicts between the network and other social constituencies. Rather, it operates within the parameters set by long-term policy objectives.

Change and Resilience in Models of Governance

The three governance models discussed in this entry are manifestations of social norms and values related to individual versus collective action, state versus market, and proactive versus reactive policies. They become institutionalized in the system of government so that state-centric governance draws on extensive institutional capabilities and a strong legitimacy of political encroachment of society, or in liberal-democratic governance where there is a search for a balance between political capabilities and individual and societal freedom.

These norms and values are reinterpreted over time. What constituted a free market in the early 20th century was different from what is today considered a free market. The developmental state was the Japanese state's response to a situation of national urgency. It became deeply institutionalized in the Japanese society for decades but was weakened during the 1980s when the downsides to aggressive progrowth policies surfaced and the most acute problems had been resolved. The Scandinavian welfare state enjoyed massive political support for several decades but has lost some of that support along with growing social affluence and increasing tax levels. The Great Society in the United States had a shorter life span and was put to rest with the election of Ronald Reagan as president in 1980.

Thus, although governance models are, by definition, institutionalized, they change over time. It is a major task in itself to account for what drives those changes. Given their institutionalization, redefining the role of the state or the market is a slow and incremental process. One type of change is where values and norms and expectations on the state are not fundamentally reassessed but instead rearticulated and given a new meaning in a new political and economic landscape. In other cases, new governance models emerge to solve new types of societal problems. Network governance could be seen as a response to growing social complexity and a lacking capacity or knowledge within the state to effectively and appropriately address those problems. If the state cannot solve such problems, actors closer to the source of problems may spring into action themselves.

Another driver of change is international influences and the diffusion of strategies and instruments for problem solving. Globalization has brought with it seamless and instant communication spreading ideas and expertise across national and jurisdictional boundaries. This means that models of problem solving, including the best agents of those measures, are imported and implemented in another national or local context, something that in turn triggers a reassessment of which actors are best equipped for such action.

Conclusion

We should not expect broad definitions of models of governance to provide clues about policy choices; as any textbook in public policy will testify, policy evolves through complex and contextualized processes. More important than promoting specific policy choices, it appears as if the governance models are manifestations of previous choices of policy styles and the role of government in solving collective problems. State-centric governance does not, in and of itself, define any long-term policy objectives. However, it does define a particular policy style. Also, the governance models define different institutional arrangements in the state, which in turn define the preconditions for broader policy choices. It is, to give an example, difficult to think of extensive welfare state policies managed overwhelmingly through network governance arrangements.

This brings us back to a comment made earlier in this entry about causal mechanisms. Do overarching policy choices such as welfare states or developmental states lead to the emergence of a particular model of governance, or is it rather the case that the entrenched governance model that is typical to a country is conducive to a particular set of broad policy goals? A tentative and preliminary answer to these questions is that there could well be causality in both directions, that is, processes of mutual reinforcement of values and norms through which one fosters the other. Megapolicy choices such as between a developmental state and a welfare state must by necessity include the creation of an institutional system equipped to coordinate and implement such large political campaigns.

Equally important, for such projects to be successful, they must draw on political and social sentiments of the significance of such projects to the country. There will only be a welfare state if dominant social constituencies support such a massive undertaking in terms of government effort and treasury, just as a developmental state requires all political, administrative, and corporate actors to commit themselves to such a national priority. Such commitment is reflective of fundamental social values such as national security or, in the case of the welfare state, social justice. Thus, social norms are directly or indirectly instrumental in fostering both overarching policy choices and predominant governance models.

The analysis presented here represents the first few steps in what appears to be a rather long journey toward a better understanding of why and how different national contexts or what Castles called "families of nations" develop different models of governance. This entry explored only a few conceivable explanatory factors. It is likely that state traditions, the trajectory of state building, political culture, and national prosperity provide important additional clues to answering those questions.

Jon Pierre
University of Göteborg
Göteborg, Sweden

See also Institutions and Institutionalism; State

Further Readings

Bell, S., & Hindmoor, A. (2009). *Rethinking governance: The centrality of the state in modern society*. Cambridge, UK: Cambridge University Press.

Castles, F. (Ed.). (1993). *Families of nations: Patterns of public policy in Western democracies*. Aldershot, UK: Dartmouth.

Esping-Andersen, G. (1990). *The three worlds of welfare capitalism*. Princeton, NJ: Princeton University Press.

Hall, P. A., & Soskice, D. (Eds.). (2001). *Varieties of capitalism: The institutional foundation of comparative advantage*. Oxford, UK: Oxford University Press.

Johnson, C. (1982). *MITI and the Japanese miracle*. Stanford, CA: Stanford University Press.

Kenworthy, L. (1995). *In search of national economic success*. Thousand Oaks, CA: Sage.

Madrick, J. (2009). *The case for big government*. Princeton, NJ: Princeton University Press.

Migdal, J. S. (2001). *State in society*. Cambridge, UK: Cambridge University Press.

Moran, M. (2002). Review article: Understanding the regulatory state. *British Journal of Political Science, 32,* 391–413.

Norton, A. (1994). *International handbook of local and regional government: A comparative analysis of advanced democracies*. Cheltenham, UK: Edward Elgar.

Peters, B. G., Doughtie, J. C., & McCullogh, M. K. (1977). Types of democratic systems and types of public policy. *Comparative Politics, 9,* 327–355.

Pierre, J., & Peters, B. G. (2005). *Governing complex societies: Trajectories and scenarios*. Basingstoke, UK: Palgrave.

Rhodes, R. A. W. (1997). *Understanding governance: Policy networks, governance, reflexivity and accountability*. Buckingham, UK: Open University Press.

Sörensen, E., & Torfing, J. (Eds.). (2007). *Theories of democratic network governance*. Basingstoke, UK: Palgrave.

Vogel, S. K. (1996). *Freer markets, more rules: Regulatory reform in advanced industrial countries*. Ithaca, NY: Cornell University Press.

Governance, Administration Policies

Some decades ago, in the early 1950s, Harold Lasswell defined politics as "who gets what." This simple but powerful definition emphasizes the fundamental point that virtually all political action is directed toward using the power of the public sector to allocate goods and services among the members of a society. Some politics is symbolic and emotive, but to some extent even those components of political action may help legitimate the political system and the decisions it must make about allocations of costs and benefits among citizens. Citizens may be willing to accept allocations with which they do not agree in part because of their emotional attachment to their government and the legitimacy associated with that attachment. This entry examines basic questions about making and implementing public policy in contemporary political systems.

Perhaps the broadest possible manner in which to consider this process of allocating goods and services through the public sector is to employ the concept of *governance*. As explained more fully below, the meaning of governance is contested rather vigorously in the academic community, but the fundamental notion of governance is that there is a need for some form of collective decision making to steer the society. This functionalist conception of governance emphasizes that although the market may be an efficient mechanism for making individual choices, there is a need to make some decisions for the society as a whole. Furthermore, those decisions may need to take into account a range of human values, not just economic values, and also attempt to accommodate influences from a range of actors. (*Value* refers to everything that is considered to have a value, be it material, ideological, ethical, or other, by most people/citizens.)

The idea of *governance* therefore is to provide some means of collective steering for the economy and society, and the word is derived from a Greek word for steering that also is the root for cybernetics. This steering activity is usually associated with government and with the formal institutions of the public sector, but this entry presents an argument for seeing the concept as neutral about the manner in which steering occurs, as suggested by Jon Pierre and Guy Peters in 2001. Although those formal institutions of the public sector continue to have a role in governing, social actors——both market and societal—also have an influence. This involvement of social actors certainly occurs in democratic societies but may also operate in less democratic systems that have sought to co-opt those social actors as part of their legitimation strategies.

The activity that we generally conceptualize as public policy making is a major component of governance. Governments exercise that steering function mainly by making policies and then putting them into action. The policy choices made through the state, and also involving actors from the civil society, are directed at a range of targets, involve the use of a number of instruments, and reflect a variety of political, economic, and social values. Furthermore, the process through which governments make policy (and again this typically is done with the involvement of other actors) is complex, often taking years to complete, and involves a large number of actors, many of whom may be able to exercise a virtual veto over the actions as suggested by George Tsebelis in 2002.

The emphasis on who gets what from political action is sometimes considered to be technical and apolitical, focusing on economic values of efficiency, or on the legalistic application of the law through public administration. Public policy is, however, anything but apolitical and reflects the convergence of a wide array of political forces, all attempting to influence the policy choices in ways that will benefit them and their constituents. According to Margaret Hill in 2005, these political forces play themselves out through complex processes that move from recognizing policy problems to deciding what good policy would be and then to evaluating the consequences of public action.

Although it is only one component of the policy-making process, governance depends heavily on the implementation of policies through a variety of political and social institutions, notably the public bureaucracy. Although much research in contemporary political science is directed at organizations and actors involved in making inputs into the political system—through voting, political parties, policy networks, legislatures, and so on—the public bureaucracy remains the largest component of the public sector, and arguably the most significant for policy choices. Not only does the expertise of the public bureaucracy inform policy making, but the bureaucracy is, in fact, more often in touch with citizens than are other parts of the public sector.

Following the implicit outline already presented, this entry begins with the most general concept of governance, moves on to the policy-making process as a major component of public action in governing, and concludes by discussing public administration and its importance in implementing, as well as making, public policy. All these aspects of governing are necessarily linked in the "real world" and in theory, but each of them also must be considered separately to understand better how the three interact. The three perspectives all combine to provide a picture of what the public sector actually does and how it affects the lives of citizens.

Studying these three aspects of political science involves the relationships between policy and administrative issues that occur in the reality of public governance and in the analytic and theoretical approaches that social scientists employ to interpret and understand those problems. Although the objective problems in these areas have considerable practical relevance, the role of theory must also be emphasized. This emphasis on theory is especially important given that some of the same strands of theory, for example, institutionalism and bounded rationality, can be applied across these areas. Therefore, the entry is a tour d'horizon designed to identify the principal issues and approaches in these parts of the discipline of political science.

Governance and Political Science

The basic idea of governance has been part of political science as long as there has been such a discipline, but the concept has become more central to studying politics over the past several decades. The exact date and source of the revival of the term is perhaps not important. What is more important is that the discipline has been supplementing its general concern with individual political behavior with greater concern with what the political processes and all that individual behavior produce by working through institutions and processes. The outcomes of governance can be conceptualized to include not only policy choices but also the general management of the political and social systems. These outcomes would therefore include the capacity to involve members of the public in the governance process, the building of trust and social capital, and what Canadians might recognize as the maintenance of "peace, order and good government," as identified in the Constitution Act of 1867.

At its most basic level, governance means the capacity to steer the economy and society toward collective goals. That definition does not privilege any particular actors or any particular institutions, but, instead, it recognizes the functional need to provide that steering and then attempts to understand how the process is carried out. One of the most important distinguishing features between governance and other forms of social action, for example, markets, is the notion of collective goals and their centrality in steering. Other forms of social action, for example, markets, assume that individual goals are sufficient, or that goals emerge through interaction, but a governance perspective assumes that goals emerge through collective choice of some sort.

Governance models in political science therefore are at their heart functionalist explanations for the processes of governing writ large. In this way, the governance approach to political science has some affinity with the structural-functionalist approaches used in comparative politics during the 1960s and 1970s. In those approaches, there was an assumption that some form of collective choice is necessary, but the specific actors to be involved in making that choice were not identified, and the variations in the way in which decisions were made were the loci for comparison. The same assumption concerning the need for collective choice is made about governance, and the approach, in the view of some political scientists, should be agnostic about how it is performed and about what activity should be subject to collective control and what is better left to the market and society.

The major contrast concerning governance that has been made in contemporary political science is between *government* and governance. The traditional assumption in political science was that governance almost inherently was the appropriate province of the institutions within the formal public sector. To the extent that nongovernmental actors were involved in the process, they were to different degrees suspect. The degree of suspicion concerning nongovernmental actors varied with political cultures. For example, the pluralist traditions in Anglo-American countries have tended to assume that interest groups pursue "the private use of public power." On the other hand, however, most of the countries of Northern and continental Europe have some tradition of corporatist involvement of private sector actors in making and implementing policy, and these countries tend not to denigrate but often laud that involvement.

This contrast between government and governance (see, e.g., Seppo Tiihonen, 2004) has emerged as a central analytic question for scholars concerned with governance. Many scholars advocating the governance perspective have argued that the formal institutions of the public sector have become increasingly incapable of providing that scarce commodity called governance. The formal organizations of the public sector are characterized as bureaucratic, clumsy, and incapable of coping with the rapidly changing societies with which they are confronted. There is ample evidence of policy failure (for a variety of evidence, see Mark Bovens, Paul 't Hart, & Guy Peters, 2001) and the underperformance of the public sector in many countries that has been used to justify the position that governments are not able to provide governance.

Critics of conventional forms of government have argued that as well as being weak in terms of efficiency, they are also weak in democratic terms. The institutions of representative democracy appear unable to capture the enthusiasm of many citizens, and levels of voting in most industrial democracies have been decreasing. Likewise, membership in political parties and many other political organizations appears to be dropping rather inexorably. Much of the successful political mobilization that has occurred has been around "flash parties" that are developed around an individual or around a particular political issue. The advocates of the "new governance," therefore, believe that providing other loci and methods of political participation will make public action more democratic as well as more capable of solving collective problems.

Even if citizens were engaged in the mass politics of representative democracy, the limited capacity of those institutions to translate their preferences into effective policy choices is clear. Richard Rose pointed out in the mid-1970s that even in majoritarian Westminster systems, voting was an extremely blunt instrument to use to control policy choices. In addition, the divided-government literature helps explain the extent to which coalition governments further complicate democratic control over policy. These constraints on governance become even more apparent as governments move away from direct ministerial mechanisms for service delivery that are driven more by market and social actors than by political controls.

The lack of agreement about what the term *governance* actually means has been addressed in part by adding various adjectives to the basic concept. One of the most important of these more circumscribed terms has been *network governance* (Eva Sørenson & Jacob Torfing, 2007), and indeed a good deal of the movement toward governance as an approach to political life has been focused on the capacity of networks to provide governance. The basic conviction of network governance is that self-organizing networks of social actors are better suited to coping with the complexity of contemporary governing demands than are hierarchical mechanisms. The "new governance" of networks is assumed to be better able to cope with that complexity and respond to changing social conditions.

Another way of limiting the meaning of governance has been to focus on "good governance," generally meaning the reduction of corruption and the creation of greater accountability, especially in less developed political systems. This version of governance has been dominant in much of the international donor community that is seeking not only to describe governance but also to improve public sector performance. Some scholars, however, have begun to focus on the quality of governance in a more general vein, attempting to understand what features of political systems may be associated with the capacity to improve the quality of life of citizens.

The linkages between institutional and behavioral characteristics of political systems and the quality of public decisions is often tenuous, but establishing that linkage can both help develop understanding governance within comparative politics and enhance the actual performance of the public sector. There is as yet very little discussion of the design of governance systems, but the accumulation of evidence may provide that version of "constitutional engineering" just as knowledge about formal public sector institutions has generated some design principles.

Governance has been discussed primarily at the national level, but the concept can be used to discuss steering capacity at all levels. For example, "global governance" has been used to describe steering at the international level. At this level, the state often ceases to be the major actor in governance, much as national governments cease to be the principal actors in some models of governance. Instead, regimes composed of nongovernmental actors are often central to providing some control over policy areas that governments may not be able to or want to. Although states have attempted to exert some control over areas such as climate control, the nongovernmental actors may prove to be more effective.

Urban areas, as well as other subnational levels, also require governance, and these regions experience all the problems of mobilizing consent and exercising control over policy choices that the national level may encounter. Indeed, networks and other forms of social actors may be more significant at the local level than at the national level, and much of the study of local politics has revolved around these interactions between formal and informal actors. Perhaps more than at other levels of government, urban areas have developed complex patterns of service delivery that correspond to the governance ideas of involving a range of private sector actors in the process of steering.

As Ian Bache and Matthew Flinders noted in 2004, one of the more important aspects of the development of governance as a means of understanding comparative politics has been the development of *multilevel governance*. While students of federalism, such as Thomas Heuglein and Alan Fenna (2006), and of intergovernmental relations more generally might consider that this term has to some extent reinvented a very old wheel, the prominence of the concept of multilevel governance has emphasized the complexity of the processes through which public programs are made and implemented. The basic point of multilevel governance is that most attempts to govern in the early 21st century involve interactions among multiple levels of government, as well as interactions between the public and private sectors.

The multilevel governance approach also points to the extent to which the contemporary state has been disaggregated and the linkage between the structure of states and the governance being provided. Christopher Pollitt and Colin Talbot noted in 2004 that as the public sector has moved a large number of its activities to agencies and quangos or has developed mechanisms such as public–private partnerships to deliver services more efficiently, the capacity for control from the center is reduced. Liesbet Hooghe and Gary Marks in 2003 pointed to the variable consequences of the different mechanisms for disaggregating the central state, but the general consequence has been to reduce the steering capacity of the center of government.

The need to understand reactions to disaggregation and the empowering of nonstate actors in governance has led to consideration of the capacity to restore some direction from the center. The concept of "metagovernance" has been coined to reflect the need for central steering the context of a more decentralized and devolved world of political action. Metagovernance is the governance of governance, or the attempt to provide or create some coherence out of very decentered policy processes. Metagovernance does not imply a return to the command and control style of governing associated with the hierarchical state but rather reflects an attempt to steer through frames and guidelines instead of through hierarchy and formal rules.

Metagovernance also represents an attempt on the part of many political leaders to reassert the primacy of politics. If networks and other forms of devolved political action are increasingly dominating political life, then the connection between voting and policy choices is becoming more attenuated. Political leaders therefore have felt the need to control policy without necessarily undermining the efficiency and legitimation gains achieved by using the devolved and marketized methods of governing. In some cases, these attempts at control have involved overt politicization of the public service, but others have involved softer forms of steering and control.

Even if one accepts that a great deal of the action of governing has been shifted away from formal institutions in the public sector and toward devolved and delegated action, it is important to remember that this activity is always being done in the "shadow of hierarchy," as noted by Fritz Scharpf (1997). That is, if the market or social actors being given grants of power fail to perform adequately in making and implementing policy, then the state can always revoke its delegation and resume control over the policy. The reassertion of hierarchical control would go beyond the softer controls associated with metagovernance to recapture direct control of policy and a more direct imposition of the priorities of elected politicians.

At the same time that governing is carried on in the shadow of hierarchy, to some extent, more formal patterns of governing are carried on in the "shadow of governance." The policy-making processes of the public sector do fail all too often; creating policy arrangements that rely more heavily on nonstate actors provide political leaders an opportunity to correct for, or perhaps avoid, those failures. In part, the capacity to move activities that may fail within the state out of the center is an important means of blame avoidance for politicians. Using market and nonstate actors can, of course, be a less "political" strategy and more a reflection of the capacity of those institutions to act in ways that the public sector cannot—whether that is being less bureaucratic in the pejorative sense of the term, being able to respond more nimbly to changing conditions, or being able to leverage human and financial resources that are out of the reach of government.

In summary, the governance turn in political science has brought the discipline back to some of its roots in thinking about the capacity of the public sector to provide coherent direction to society. Clearly, in the case of failed states, this type of governance is not possible, and even for some more successful political systems, it can be sporadic and inconsistent. Furthermore, governance is not just about the success or failure of the formal public sector but also involves the capacity of those institutions to interact successfully with social actors and also increasingly the capacity to interact with the global environment.

Public Policy

Public policy studies are the second major body of work to be considered when thinking about the outputs of the public sector. Most political scientists do not think of themselves as "policy scientists" of the type advocated by Harold Lasswell and Abraham Kaplan in 1950, but in essence they are. The individualistic turn in political science, reflected in both behavioral and rational choice approaches to the discipline, has diverted attention from the output side of the political system toward aspects of political life where institutions and processes are less dominant features. For many political scientists taking those approaches, public policy is better left to economists or to professions such as planning.

Although the two aspects of political science—policy and individual behavior—are in fact more complementary than competitive, they tend to ignore each other and to press the primacy of one or the other. This is most unfortunate, given, that it is individuals, working singly or collectively, who must make policy decisions; we need to know more about how they make those decisions. Likewise, public policies frame a good deal of the political action by individuals, so that voting may be voting for or against specific policy choices.

The Policy Process

The study of public policy in political science has several important but somewhat disparate dimensions. The largest single body of research has been on the policy process, notably the work of Charles O. Jones. This approach to policy is inherently political, arguing that the policy choices made by government are primarily a function of the political process through which they are made. The majority of the process models attempt to explain policy choices by understanding the actors who are involved at each of these stages, as well as by understanding linkages among the stages. These models are, however, far more useful as heuristics for describing the process than they are as means of explaining outcomes.

The conventional process model of policy making begins with agenda setting or problem definition and then proceeds through a series of steps such as program design, legitimation, resource attachment, and evaluation; typically, it did not have any strong assumptions about the political mechanisms that were used to manage that movement. Although some elements within the general model, for example, agenda setting, relied on the role of policy entrepreneurs to provide agency within the process, the general models tended to be largely devoid of that animation.

Furthermore, the process model as generally used in political science did not reflect the degree of conflict that might exist in these crucial political processes; rather, the process appeared to move along rather smoothly to its conclusion. One important addition to process models that did contain that conflict was Paul Sabatier and Hank Jenkins-Smith's (1993) "advocacy–coalition framework." This model was concerned explicitly with policy change and suggested that policy change came about through the clash of ideas, as manifested in an existing policy coalition and a potential replacement. These conflicts could be worked out through bargaining processes, often generating a new synthesis about policy, which would in turn be institutionalized, setting the stage for the next round of policy change.

In addition to the general model of the policy process, there has been a substantial degree of theorizing about various components of the policy process. In particular, political scientists have been concerned with agenda setting, budgeting, implementation, and, at one point, evaluation. These process considerations have, however, often been divorced from concerns about the quality of policies being adopted. Each of these process areas has produced a large body of literature that has influenced more general studies of policy. That having been said, with the exception of the study of agenda setting, these areas of research have declined in political science to some extent, reflecting the general decline in the importance of policy studies in the discipline.

Implementation

The first of these components of the policy-making process to emerge as a distinctive field for research was *implementation*. Although the term is hardly new, it was brought to center stage in policy studies by Jeffrey Pressman and Aaron Wildavsky in 1974. They focused attention on the difficulties of taking legislation and making it work in "the real world" and especially the difficulties in making it work in the manner intended by the framers of the legislation. They developed this analysis by identifying the number of clearance points that had to be passed successfully if the program were to go into effect. While their analysis of the difficulties of implementation was perhaps excessively pessimistic, it did help frame a fundamental policy problem in a strong analytic model.

The Pressman and Wildavsky framework for understanding implementation is a top-down perspective on the process, while the process may also be analyzed from a bottom-up perspective. The bottom-up perspective assumes that programs might be better designed from the perspective of the prospective of implementation. This "backward-mapping" perspective assumes that programs might ultimately be more effective if their final implementation were central to their formation. Although criticized as putting implementation ahead of policy goals, this approach does emphasize the key role of implementation in making programs actually work.

A third wave of implementation research emphasized quantitative studies of the success and failure of implementation efforts, relying heavily on measures that were in many ways difficult to sustain in the face of the complexity of the implementation process. If anything, the complexity of implementation has become even greater. For example, implementation increasingly involves multiple actors at multiple levels of government, as described by Laurence O'Toole in 2004, or perhaps also actors in the private sector.

Budgeting

Wildavsky was also central to the emergence of public budgeting as an area of concerted research in political science in the mid-1960s. In contrast to economists who looked at the public budget as a search for a Pigovian optimal allocation, Wildavsky saw budgeting in a much more political manner, emphasizing the importance of structure and process in producing the outcomes. In particular, the repetitive nature of budgeting and the multiple organizations involved in the process (at least in the United States) tended to produce incremental outcomes. There was also some individual element involved in this model of incrementalism, given that many of the actors involved played the "game" year after year and their memories and their interpersonal trust also influenced the outcomes.

After his initial insights into the American budgetary process, Wildavsky and his colleagues, Otto Davis and Michael Dempster, made a number of other important contributions to a political understanding of budgeting. First, they formalized and tested the incremental model of budgeting in 1966 and demonstrated that there was a great deal of predictability in the decisions made for the U.S. federal budget. Furthermore, Wildavsky and colleagues developed alternative models that could describe budgetary choices in other types of political systems. Finally, Hugh Heclo and Wildavsky in 1974 provided a detailed analysis of budgetary decision making in the British Treasury, again emphasizing the importance of trust in these deliberations.

The seminal work by Wildavsky sparked a number of interesting elaborations of budgeting and critiques of the incremental model. Unfortunately, however, that interest in budgeting has not been maintained, even given its centrality in governing. The interest in budgeting that remains has been largely on the part of rational choice scholars who see the budget as a logic locus for actors to seek to maximize their own benefits and to pursue interesting strategies in order to do so.

As Charles Lindblom (1965) notes, the incremental ideas that have been central to understanding budgeting in the United States, and to a lesser extent elsewhere, are related to more general analytic approaches emphasizing the likelihood, and the desirability, of incremental decision making. In this view, the public sector will produce better governance through successive limited comparisons rather than through attempting to make a single, correct, and comprehensive decision. While such a pattern of governing can be condemned as excessively conservative, the logic also corresponded well to the American political style from which it had evolved.

At a more theoretical level, beginning with Herbert Simon's (1947) work, the logic of *bounded rationality* has been a powerful contender in the attempted domination of more rationalistic