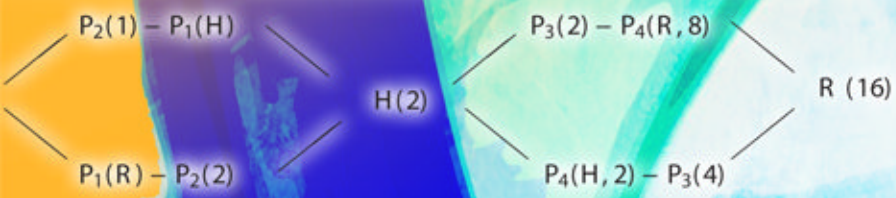
Russell J. Jenkins
Walter E. Sullivan
Editors

PHILOSOPHY OF MIND



WORLD PHILOSOPHY SERIES

NOVA

WORLD PHILOSOPHY SERIES

PHILOSOPHY OF MIND

WORLD PHILOSOPHY SERIES

Additional books in this series can be found on Nova's website under the Series tab.

Additional e-books in this series can be found on Nova's website under the e-books tab.

WORLD PHILOSOPHY SERIES

PHILOSOPHY OF MIND

RUSSELL J. JENKINS
AND
WALTER E. SULLIVAN
EDITORS

nova publishers
New York

Library of Congress Cataloging-in-Publication Data

Philosophy of mind / editors, Russell J. Jenkins and Walter E. Sullivan.
p. cm.
Includes index.
ISBN: 978-1-62257-229-8 (eBook)
1. Philosophy of mind. I. Jenkins, Russell J. II. Sullivan, Walter E.
BD418.3.P4477 2012
128'.2--dc23

2012021147

Published by Nova Science Publishers, Inc. † New York

CONTENTS

PREFACE

In this book, the authors present current research in the study of the philosophy of the mind. Topics discussed in this compilation include the concepts of hope and belief; how consciousness builds the subject through relating and human behavior; analyzing the neurophysiological mechanism of qigong on the mind and brain activity; the conscious and unconscious mind and implications for society, religion, and disease; how the mind is shaped by culture; and the power of computational mathematics to explore some of the universal ways by which each human mind builds its image of the world.

Chapter 1 - Although the concept of "hope" received attention from such notable philosophers as Aristotle, Aquinas, Spinoza, Hume and Kant, and is much discussed in religious philosophy and psychology, it has not been subjected to much analysis by modern analytical philosophers. There is general agreement that "A hopes that Q" is to be analysed as "A desires that Q; and A is uncertain that Q (or considers Q to be possible)". There is, however, disagreement about the sense of uncertainty (or possibility) involved. By far the most comprehensive modern philosophical analysis of "hope" is provided by John Patrick Day, who interprets "A is uncertain that Q" as "A attaches a subjective probability to Q of more than 0 but less than 1". Day also holds that "A hopes that Q" is compatible with "A believes that Q", because he maintains that "A believes that Q" is to be analysed as "A attaches a subjective probability to Q of more than ½ but less than 1". In this Chapter, it is argued, with particular attention to Day's analysis, that "A hopes that Q" is better analysed as "A desires that Q; and, with Q in mind, neither believes nor disbelieves that Q" where "A believes that Q" is to be analysed as "A affirms that Q is actually, rather than probably or merely possibly, the case". Interestingly, Kant held that hope is compatible with belief (viewed as

affirmation), though not with knowledge, being equivalent to affirmation grounded in what ought to be the case. It is argued that if Kant's notorious moral argument for God is reinterpreted as an argument for hope (analysed as precluding affirmation) that God exists, then it can withstand objections commonly brought against it.

Chapter 2 - This chapter aims to show that the main difference that consciousness makes to human behavior is to provide us with a sense of self. Consciousness does this by allowing us to relate ourselves to other entities, and therefore to understand what kinds of relations exist between us and them. Variations in the state of nervous energy elicited by the use of attention are the basic underlying mechanism of consciousness. They are used to put things in relation, mainly by acting as the basis for the construction of possible orders (such as space and time).

Chapter 3 - With the increasing demand of non-pharmacologic treatment for psychological problems, qigong has been shown to provide promising physiological and psychological effects. In view of the inadequate documentation analyzing the neurophysiological effects of qigong, this chapter aimed at analyzing the underlying neurophysiological mechanism of qigong on the mind, especially on the brain activity. The brain activity under qigong state is quite different from that under other relaxed states (e.g. sleep and close-eyed rest), however, it is rarely mentioned in previous literature. To further understand this, studies employing neuroelectrical (e.g. EEG) and neuroimaging measurements (e.g. fMRI) were extracted from seven databases for analysis. Both Chinese and English written articles were included.

Findings showed that mind regulation and breath regulation of qigong training help to stabilize the autonomic and stress response systems. The unique neurophysiological mechanism of qigong, which is different from other relaxing mindful state, is characterized by activation of the parasympathetic system, improvement in cortical-subcortical synchronization, thalamic performance, etc. The authors suggest that qigong can be a psychosomatic exercise of moderate intensity to stabilize the mind. However, due to the special nature of EEG and fMRI measurements, within-group design and relatively small sample size were commonly used in available studies. Therefore, more sophisticated RCT design with larger sample in future research is recommended for further justification.

Chapter 4 - Decisions are based on estimates of *a priori* probability, evidence and the values and costs of the anticipated outcome. An optimal decision strategy is one which seeks to maximize expected value. Digital computers, neural networks, robots, the unconscious brain and the conscious

mind are all capable of following an optimal decision strategy. The rules by which values and costs are attached to specific outcomes are pre-programmed but then modified by experience even in automata. The conscious mind, however, must experience the values and costs and therefore it is directly rewarded or punished by the unconscious brain. This is the profound insight of psycho-analysis expressed in modern parlance. Molecules circulating in the blood act directly on the brain controlling mood and motivation. They switch on and off genetic and proteomic networks that control and modulate conscious experience. The rules of human behavior influenced by evolutionary genetics but extensively modified by social experience are written in the networks and used to reward or punish the conscious mind. These ideas have implications for normal social functioning, altruism, aesthetics, the determinants of happiness and the universal phenomenon of religion. But complex genetic and proteomic networks can malfunction and lead to disabling diseases such as schizophrenia, depression and functional psychiatric disorders. Faulty genetic networks caused by deleterious mutations might play a part, particularly in the etiology of schizophrenia, but the exciting possibility that microbial molecules play a major role is also explored. If a genetic or proteomic network is switched on or off without concomitant inflammation possibilities include bacterial toxins absorbed from the intestinal tract or antibodies directed against microbes which cross react with brain proteins.

We will improve our health once we appreciate that social interaction can lead to happiness through acts of altruism which generate conscious reward. Our spiritual life is enhanced by the realization that religion will inevitably emerge when intelligent conscious animals interact. There is a good chance we can reduce disease, including diseases of the mind, by controlling the rate, dose and route of microbial exposure so as to optimize the immune response and reduce the chance of infection and the generation of auto-antibodies. All this follows from a philosophy of the mind in which information theory and decision theory play a key role. This philosophy gives us a deeper understanding of the nature of consciousness and points to practical solutions to problems that currently trouble human kind.

Chapter 5 - The claim that mind is shaped by culture – as nowadays assumed by many authors and theoretical perspectives in philosophy, psychology and social sciences – is critically discussed. Firstly, three positions concerning the contribution made by culture to the alleged construction of the mind are highlighted: (i) culture contributes to the construction of the mind by offering *opportunities* that allow the endogenous psychological resources to be made explicit; (ii) culture contributes to the construction of the mind by

producing *influences* on it; (iii) culture is the *condition* to construct the mind. The last position leads to problematic consequences which are contested. If mental experience arises by assimilating the dominant cultural framework, where do questions about the cultural framework itself come from? People sometimes realise that there is something which fails to fit the shared cultural framework. On the opposite, a cultural framework is accepted because it is realised that it gives an adequate explanation of aspects of reality that to a certain extent had already been perceived outside the framework itself. Finally, radical culturalism fails to give reason of why and how changes in the existing framework occur: if mind is shaped by current culture, where do original ideas come from? These arguments lead us to concede that there are mental experiences not mediated by culture which are the source of debating and innovation. In conclusion, the culturalist perspective reminds us that mind is culturally constrained but can not induce us to believe that it is merely a cultural construction lacking of a genuine psychological status.

Chapter 6 - Like Philosophy, Mathematics deals with abstract ideas, i.e. immaterial objects which inhabit and work in the Mind. The chapter "Computation in Mind" proposes to use the power of computational mathematics to explore some of the universal ways by which each human mind builds its "imago mundi", its image of the world. The primary focus is put on epistemology and the use of mathematics is minimal, relegating the necessary technical details to an appendix. The Chapter develops the viewpoint that Science and Mind are mirror images for each other which use specific calculations over three kinds of numbers. It presents some epistemological consequences of the lack of associativity or commutativity for the two basic operations which are × and + when the calculations are performed over vectors or over matrices. The evolutive nature of the scientific logic is illustrated on several examples. In particular induction in computation suggests that any matrix ring can be usefully considered as a structure of macroscalars.

In: Philosophy of Mind
Editors: R. J. Jenkins, W. E. Sullivan
ISBN: 978-1-62257-215-1

Chapter 1

HOPE AND BELIEF

Deryck Beyleveld
Law and Bioethics, Durham University, UK
Moral Philosophy and Applied Ethics, University of Utrecht,
The Netherlands

ABSTRACT

Although the concept of "hope" received attention from such notable philosophers as Aristotle, Aquinas, Spinoza, Hume and Kant, and is much discussed in religious philosophy and psychology, it has not been subjected to much analysis by modern analytical philosophers. There is general agreement that "A hopes that Q" is to be analysed as "A desires that Q; and A is uncertain that Q (or considers Q to be possible)". There is, however, disagreement about the sense of uncertainty (or possibility) involved. By far the most comprehensive modern philosophical analysis of "hope" is provided by John Patrick Day, who interprets "A is uncertain that Q" as "A attaches a subjective probability to Q of more than 0 but less than 1". Day also holds that "A hopes that Q" is compatible with "A believes that Q", because he maintains that "A believes that Q" is to be analysed as "A attaches a subjective probability to Q of more than ½ but less than 1". In this Chapter, it is argued, with particular attention to Day's analysis, that "A hopes that Q" is better analysed as "A desires that Q; and, with Q in mind, neither believes nor disbelieves that Q" where "A believes that Q" is to be analysed as "A affirms that Q is actually, rather than probably or merely possibly, the case". Interestingly, Kant held that hope is compatible with belief (viewed as affirmation), though not with

knowledge, being equivalent to affirmation grounded in what ought to be the case. It is argued that if Kant's notorious moral argument for God is reinterpreted as an argument for hope (analysed as precluding affirmation) that God exists, then it can withstand objections commonly brought against it.

INTRODUCTION

The nature of hope has been a topic for philosophical reflection at least since the ancient Greeks and Romans. Aristotle [2, 449b] has views about it. So, too, do Aquinas [1], Spinoza [26, Part III, Definition of the Emotions], Hume [14, Book II Part III Section IX], John Stuart Mill [16], and Kant. Indeed, for Kant:

> All the interests of my reason, speculative as well as practical, combine in the three following questions:
>
> 1. What can I know?
> 2. What ought I to do?
> 3. What may I hope? [18, A805 B833]

Despite this, the concept of hope has not received a great deal of attention by modern analytical philosophers, although it is much discussed in the philosophy of religion and by psychologists (e.g., [15], [27]).

It is widely accepted that "A hopes that Q"$\leftrightarrow$[1] "A desires Q; and A is uncertain (in *some* sense) that Q". The conative condition, "A desires that Q", covers "A wishes, wants, or values Q", indeed any positive attitude A has to Q, while "A does not desire that Q" covers any case where A wishes, or wants to avoid Q, or otherwise finds Q unattractive or unacceptable. The intensity of A's hope for Q will vary with the intensity of A's desire for Q. "A fears that Q" $\leftrightarrow$ "A does not desire that Q; and A is uncertain that Q". "A hopes that Q" $\leftrightarrow$ "A fears that $\sim$Q".[2] If A neither desires nor does not desire that Q (i.e., A is conatively indifferent to Q), A neither hopes nor fears that Q.

---

[1] "Entails and is entailed by" or "mutually entails".

[2] Spinoza (see [26, Part III, Definition of the Emotions XIII]) and Hume (see [14, Book II Part III Section IX]), both emphasise that hoping always involves fearing and vice versa. In ordinary use, "A fears Q" is ambiguous. It can mean, "A does not desire Q" $\equiv$ "A fears_1Q" (e.g. "Paul fears spiders, is frightened of them", or it can mean "A fears_2 Q" $\equiv$ "A fears_1Q;

My first aim in this chapter is to present an analysis of the cognitive condition, of the sense in which hope involves uncertainty about its object. I will argue that hoping that Q involves, *with the question whether Q is or is not the case in mind,* neither believing nor disbelieving that Q. I consider that a sound account of the idea of hope is vital for a sound account of rational action generally. I also consider that beings are subjects and objects of moral concern essentially because they are hoping and fearing subjects. These topics are too broad to pursue further here.[3] Here I will restrict application of my analysis to reflection on Kant's moral argument for the existence of God. I will argue that this argument is valid if it is understood not as an argument for believing that God exists on the presumption of a categorical imperative, but as an argument for hoping (in the sense I advocate) that God exists.

To place my analysis of hope in context, and to explain clearly how my arguments will proceed, it is necessary to begin with an outline of the different positions taken on the cognitive position.

Positions on the Cognitive Condition

The definition of hope provided by St Thomas Aquinas provides a useful starting point. According to Aquinas, hope is "a movement of appetite aroused by the perception of what is agreeable, future, arduous, and possible of attainment" [1, p.7]. If we relate this to the widely accepted idea, Aquinas holds that "A hopes that Q" $\leftrightarrow$ "A desires that Q"; and "A is uncertain that Q (for which the necessary and sufficient conditions are: A perceives Q to be future, arduous, and possible of attainment)".

"Q is future" means "Q is not already the case". If A must perceive Q to be *objectively* future, then this disqualifies statements like, "Peter hopes that Liverpool won their match against Arsenal last night", "Sally hopes that it is not raining in Chicago now", and "David hopes that he has won the race he

and A is uncertain that Q" (e.g., "Paul fears that the spider on the carpet will bite him"). So "A fears_1 Q" means "Q is the object of A's fear_2 that Q" and "A fears_2 Q" is the contrary of "A hopes that Q". Ambiguity is removed if "A fears Q" is taken to mean "A does not desire that Q" and "A fears that Q" is taken to mean "A does not desire that Q; and A is uncertain that Q".

[3] The first of these topics is important for my project on the role that precautionary reasoning can legitimately play in law (on which see [6]). The second claim is an implication of the analysis of dignity in [5] (esp. Chapter 5), which also contains less refined reflections on Kant's moral argument for God than I present here. See also [4].

has just finished running," none of which have objectively future objects. Q, or ~Q, has already happened or is happening.

John Patrick Day says that Q must be *subjectively* future, not *objectively* future, "that which is not yet within the subject's experience" [8, p.23] in the sense that the subject "does not yet know of it" [8, n.3, p.28]. This, I think, is a key insight.

However, "to know of it" is open to interpretation. Since it is, in *some* sense, to be "certain" of it, we may say that when Q is subjectively in the future for A, A's current *perception* of Q is, in some sense, uncertain.[4]

"Q is arduous" implies that Q is not already given, inevitable, pre-determined, or unavoidable: something must *be done* to achieve Q. However, since objects of hope are not necessarily ends of action (e.g., A might hope that the sun will shine, or that God exists), if "Q is arduous" is to be a universal condition, it must mean that *the existence* of Q (regardless of whether or not this requires action) is not certain, or that it is not necessarily the case that Q.[5] So, "Q is not certain" must be taken to mutually entail "Q might not be the case" ≡ "*~Q is possible"*.

Conversely, "Q is possible" mutually entails "~Q is not certain". Therefore, we may state the combined condition in any of the following equivalent ways: "A perceives that neither Q nor ~Q is certain or impossible"; "A perceives that neither Q nor ~Q is certain"; "A perceives both Q and ~Q to be possible"; "A perceives Q (or ~Q) to be neither impossible nor certain", which I will formulate as "Asπ(Q) ≠ 0 and≠ 1".

Hence, Aquinas' definition may be restated as: "A hopes that Q" ↔ "A desires Q; and A is uncertain that Q (for which the necessary and sufficient conditions are: A's current perception that Q is the case is uncertain; and Asπ(Q) ≠ 0 and≠ 1". This will be formulated as: "A hopes that Q" ↔ "A desires that Q; and Asu(Q) = >0<1 (↔Asf(Q) and [Asπ(Q) ≠ 0 and≠ 1])".

"Asu(Q) = 1" means "A is certain that ~Q", whereas "Asu(Q) = 0" means "A is certain that Q". "Asf(Q)" means "Q is subjectively future for A" (i.e., "A's current perception of Q is uncertain"). "Asπ(Q) = 0" means "A perceives that Q is impossible (~Q is certain)". "Asπ(Q) = 1" means "A perceives that ~Q is impossible (Q is certain)".

---

[4] Compare Spinoza: "Hope is an inconstant pleasure, arising from the idea of something past and future, whereof we to a certain extent doubt the issue" [26, Part III, Definition of the Emotions XII]).

[5] "That Q" or "Q is the case" will generally to be taken to cover objectively present, future and past objects.

The modern orthodoxy is simply: "A hopes that Q" ↔ "A desires that Q; and "Asπ(Q) ≠ 0 and ≠ 1".[6] Day, who presents by far the most sustained modern analysis of the concept, rejects both this and Aquinas' definition, because, following David Hume [14, Book II Part III Section IX], he maintains that "Hope and Fear must be analysed in terms of Subjective Probability and not in terms of Subjective Possibility" [8, p.24]. This is because there are degrees of Hope and Fear and also degrees of Subjective Probability, but no degrees of Subjective Possibility. We say, *e.g.* (1) 'A has high hope that Q', [and] (2) 'A has only a faint hope that Q'… [which] … it is impossible to analyse … in terms of Subjective Possibility, as the reader can verify for himself [8, pp.24-25].

Therefore he gives the following definition: "A hopes that Q" ↔ "A desires (in some degree)[7] that Q; and A believes the subjective probability[8] of Q =>0<1" (See [8, p.19; p.25]].[9]

Day's definition will be formulated as: "A hopes that Q" ↔ "A desires that Q; and Asp(Q) = >0<1". "Asp(Q) = >0<1" replaces "Asf(Q)" *and* "Asπ(Q) ≠ 0 and≠ 1" in my reconstruction of Aquinas' definition. So what has happened to "Asf(Q)"? It seems that Day considers that "Asu(Q) = >0<1", is to be analysed as "Asp(Q) = >0<1", which (because of Asf(Q)) is to be interpreted as "A *currently* estimates sp(Q) as >0<1".

I propose, instead, that: "A hopes that Q" ↔ "A desires that Q; and Asb(Q) ≠ 0 and ≠ 1". "Asb(Q)" stands for "A's tendency to believe that Q" (or

---

[6] See, e.g., J. J. Godfrey [11, p. 30]; J. Harrison ("Hoping that something is so implies two things, wanting it to be so . . . and neither believing that it is inevitable . . . nor impossible" [12, p.80]); H. H. Price, "If we hope that x will happen, we must at least believe it [not just logically but causally] possible that x will happen" [25, p. 268] (and that it will be a good thing if x happens); and S. R. Sutherland [28, pp. 197], who prefers Harrison's formulation on the grounds that insisting on causal possibility rules out, e.g., hope that God exists.

[7] Day always writes this qualification in. If I leave it out it is to be understood as being there.

[8] "Probability" covers both mathematical probability and "Inductive Probability (or Degree of Confirmation)" [8, p.35; p.87], and degrees of the latter can only, strictly speaking, be "expressed by 'very probable', 'fairly probable' etc., but cannot be measured by real numbers" [8, p.35] (see also [8, p.87]). Hence, any assignment of numerical values to them is purely for convenience (see 8, p.87]. By the "subjective probability of Q", Day means "the degree of probability which the subject, A, believes (estimates, thinks, judges) the object, Q, to possess" [8, pp.87-88].

[9] See also, R. S. Downie, according to whom, rather like H. H. Price [25] logical possibility is not enough, because "the typical cases of 'hope that' fall within a narrower scale—that of physical probabilities and likelihoods. By this I mean that we can only hope where we believe that there is a positive probability that the object of hope will be realized" [9, p.249]. Day's definition is also in line with J. M. O Wheatley [30, p. 127], according to whom hoping always involves some degree of expectation. However, there is no reason to think that Downie or Wheatley would follow Day's elaboration of this definition.

"the degree to which A *leans towards* believing that Q" or "the believability of Q for A"). "Asb(Q) = 0" means "A disbelieves that Q" ≡ "A believes that ~Q". "Asb(Q) = 1" means "A believes that Q". Asb(Q) ≠ 0 and ≠ 1" means "A neither believes nor disbelieves that Q" ≡ "A unbelieves that Q".[10] "A does not believe that Q" ↔ "'Asb(Q) = 0' or 'Asb(Q) ≠ 0 and ≠ 1'".

I contend that "Asf(Q)" is to be analysed as "Asb(Q) ≠ 0 and ≠ 1". However, *provided that A has Q in mind*, "Asb(Q) ≠ 0 and ≠ 1"→ "Asπ(Q) ≠ 0 and ≠ 1" but not vice versa, and then "Asπ(Q) ≠ 0 and ≠ 1" is cognitively redundant for "A *does* hope that Q" though not for "A *could* hope that Q". However, there is no particular harm in contending that "A hopes that Q" ↔ "A desires that Q; Asπ(Q) ≠ 0 and ≠ 1; and Asb(Q) ≠ 0 and ≠ 1".

Day's opposed conception involves a view of "believing that" very different from my own. The concept of belief that I will employ, for reasons that will become clear, specifies that to believe that Q is to affirm Q, which is to be committed to treat "Q is the case" as a true proposition in thought and action. There are no degrees of believing that Q, though there can be degrees of leaning towards believing that Q, degrees of justification for believing that Q, and degrees of resistance to giving up believing that Q. But, "A believes that Q" entails nothing about the degree of justification or confirmation that A thinks there is for "Q is the case". Day, on the other hand, holds that "A believes that Q" → "Asp(Q) = >1/2<1".

> [D]egrees of belief that Q correspond to the degrees of sp(Q) in the interval (1/2,1) exclusive. E.g. "A firmly believes that Q" corresponds to "sp(Q) = 2/3". It follows [because, when A is hopeful that Q, "A estimates the probability of Q as >1/2<1" [8, p.56]] … that "A is hopeful in some degrees that Q" is equivalent to (entails and is entailed by) "A desires in some degree that Q, and A believes that Q".[11] … Belief that Q, then, entails the existence of a second-order belief about the probability of Q. [8, pp.72-73]

My argument will proceed as follows. First, I argue that the correct analysis of "Q is subjectively future for A" is "A currently neither affirms nor disaffirms that Q". Hence, if "belief that" is equated with "affirmation",

---

[10] "Asb(Q) ≠0 and ≠ 1" may also be written as "Asb(Q) = >0<1'" and read as "the subjective *probability* of Q for A = >0<1", *but only when* "A thinks that it is *probably true* that Q" ("Asb(Q) = >1/2<1") is contrasted with "A thinks that it is *actually true* that Q" ("Asb(Q) = 1") not with "A thinks that it is *certainly true* that Q" (which Day has as "Asp(Q) = 1").

[11] Day never actually says that "A believes that Q" means "Asp(Q) = >1/2<1", but it follows logically from these statements that he is committed to this view.

"Asb(Q) ≠ 0 and ≠ 1" is necessary for "A hopes that Q". I explicate belief as affirmation, and explain how it relates to various concepts of uncertainty, and defend the thesis that Asb(Q) = 0 (or = 1) is sufficient to negate hope against the obvious objection to it.

Secondly, I argue that Day's view of belief is not merely incompatible with belief as affirmation: it sets up an infinite regress that can only be avoided if beliefs are affirmations or "A believes that Q" → "Asp(Q) = 1". Either way, "A hopes that Q" cannot → "A believes that Q". Day's view over-rationalises the concept of belief. His justification for his view is inadequate and cannot deal with these objections.

Thirdly, I respond to Day's claim that "hope" cannot be analysed without reference to subjective (evidential) probability. The reason he gives for his claim has no bearing on the conditions for hope; it only bears on the degree to which the hoping subject considers it likely that the hope will be fulfilled.

I then apply the concept of hope developed to Kant's moral argument for the existence of God. I argue that if Kant's reasoning is interpreted or reconstructed as an argument for the rational/moral necessity of hoping that God exists in the terms of my analysis of "hope" rather than as an argument for the rational/moral necessity of faith (as involving belief) that God exists, then it survives the standard objections that are brought against it. This has, however, the radical consequence that to be committed to the idea of morality as categorically binding is to be committed to the idea that if God exists then God does not want us to believe that God exists.

My View

Hope and the Future

According to Day, an object of hope (or fear) is always *subjectively* future. For example:

> Jack hopes that Jill has caught her train. Jill's having caught her train is objectively past, but subjectively (*i.e.* for Jack) future" [8, p.22]. "That which is objectively ("actually") past may also be subjectively future for A, in the sense that A does not yet know of it" [8, n.3, p. 28].

However, this must not be confused with the false theory which analyses "A hopes that B has caught her train" as "A hopes that A will learn that B has

caught her train". For, here, A's learning is of course objectively future, not subjectively future [8, n.3, p.28].

This is partly sound. In order to highlight the tenses involved in Day's example, I will restate it as: "Jack hopes now that Jill yesterday caught her train".

Whether Jill caught her train or missed her train, she did so yesterday (in the objective past), and John perceives now (in the objective present) that both alternatives are objectively past events. But Jack now is "uncertain" (i.e., does not perceive) which of these events occurred yesterday. While the object of Jack's hope (Jill yesterday caught her train) and the object of his correlative fear (Jill yesterday missed her train) lie in the objective past, John's perceiving that Jill yesterday caught — or missed — her train, lies in the objective future. Hence, to say that an event is subjectively future is to say that whether the event is objectively past, present, or future, the subject has not yet come to perceive the event as being actual (objective). It lies in the objective future whether or not the subject will come to perceive the event as part of the actual past, present, or future world.

Day's statement that "Jill's having caught her train is objectively past, but subjectively (i.e. for Jack) future", and that what "is objectively ('actually') past may also be subjectively future for A, in the sense that A does not yet know of it", is not inconsistent with this. Day is also right that this must not be confused with the false idea that "A hopes that B has caught her train" means "A hopes that A will learn that B has caught her train".

However, the difference between the false theory and the correct one is not that the false theory makes it a condition of "A hopes that Q" that "A's learning that Q" is an objectively future event instead of something subjectively future. The correct theory makes both A's learning that Q and A's learning that not-Q possible objectively future events! The reason why the false theory is false is that Jack is not hoping that he will come to perceive that Jill yesterday caught her train. He is hoping that Jill yesterday caught her train.

Having made this mistake, Day tells us that "the objects of hoping and fearing must be propositions about the future" [8, p.59]. This is also false. The object of Jack's hope (Jill yesterday caught her train) is, if actual, an objectively past event, not a proposition at all, let alone a proposition about the future. The proposition "Jack now hopes that Jill yesterday caught her train" describes an objectively present event (Jack's present hoping). The only thing that lays in the future — and this is the objective future — is Jack's coming to perceive whatever he will come to perceive about what Jill (yesterday) did. While this may be expressed by the proposition "Jack has not yet perceived

that Jill yesterday did (or did not) catch her train", this is a proposition about Jack's present ambivalence concerning whether or not Jill yesterday caught her train. The only relevant thing that lies in the future is the psychological fulfilment or dashing of Jack's hope that Jill yesterday caught her train.

Of course, we can say that Jack, in hoping that Jill yesterday caught her train, hopes that the proposition "Jill caught her train yesterday" is true. And this does make the truth of this proposition the object of Jack's hope. But its truth (which does not depend on when or even if it is perceived to be true) rests on its correspondence with the actuality of Jill's yesterday catching her train; while Jack's perception of its truth (if that happens) lies in the objective future. The only way in which the objects of hoping and fearing can be propositions about the future is if "A hopes that B has caught her train" means "A hopes that A will learn that B has caught her train",[12] which is untenable.

The crucial question, however, remains: "What does it mean to say that A has yet to perceive what actually happened/is happening/or will happen?" I agree that it means that A has an uncertain perception of whether Q or ~Q is the case (has happened/is happening/will happen).

So, with our example in mind, the question becomes, "In what sense is Jack uncertain (ambivalent) as to whether Jill yesterday caught or missed her train?" "What change in Jack's mental state would constitute the removal of Jack's uncertainty (ambivalence) about whether or not Jill yesterday caught her train?"

Because Day analyses "Asu(Q) = >0<1" as "Asp(Q) = >0<1", it follows that Asf(Q) ("Q is subjectively future for A") simply places a temporal restriction on Asu(Q) = >0<1/Asp(Q) = >0<1. "Asf(Q)", "Asu(Q) = >0<1", and "Asp(Q) = >0<1" are to be treated as equivalent to "A *currently* considers neither Q nor ~Q to be conclusively confirmed: each has a probability >0<1". Only when A judges that p(Q) = 1 (or = 0) does A cease to be uncertain that Q and Q cease to be subjectively future for A. Q does not cease to be subjectively future for A when A comes to believe that Q (or that ~Q), no matter how firmly A believes that Q (or that ~Q), because "A believes that Q" → "Asp(Q) = >1/2<1". Since "A claims to know that Q" → "Asp(Q) = 1" (see [8, 72]), "A hopes that Q" → "A currently does not think that A knows that Q (or that ~Q)".

[12] Day claims that "A hopes that Q" is neither true nor false, because "it is arguable that propositions about the future are neither true nor false" [8, p.59]. There is no need to debate this here, given that it is false that "A hopes that Q" is in any relevant sense a proposition about the future.

Is this really what the ambivalence that constitutes "Asf(Q)" amounts to? Let's return to our example. What essentially characterises the object of Jack's hope, Jill having yesterday caught her train (Q) being subjectively future for Jack, is that neither the proposition "Jill yesterday caught her train" ("Q is the case") nor the proposition "Jill yesterday missed her train" ("~Q is the case") as yet describes part of Jack's mental representation of the actual world. Both propositions are potentially items in Jack's present representation of the actual world, but neither has yet become part of it. Each proposition, now, describes (for Jack) only a potential fact. Jack thinks it is possible that "Q is the case" is actually true, but also possible that "Q is the case" is not actually true. To say this is to say no less, and no more, than Jack does not as yet consider either proposition to be actually true. So, to say that the event referred to by "Jill yesterday caught (or did not catch) her train" is subjectively future (for Jack) is to say that Jack is in a state of mind that envisages both the proposition "Jill yesterday caught her train" and the proposition "Jill yesterday failed to catch her train" as only potentially true. To coin a metaphor, they are merely candidates for membership of the club that constitutes Jack's idea of the actual world. At the moment, as far as Jack is concerned, both "Q is the case" and "~Q is the case" are only possible members of Jack's actual world. But this is possibility v. actuality — "possibly true" v. "actually true"; not "possibly true" v. "certainly true".

Jack now only accepts that Jill yesterday might or might not have caught her train. As such, he does not unequivocally accept either that Jill yesterday did catch her train or that Jill yesterday did not catch her train. Whatever else he thinks about the matter (and suppose that he thinks it is more probable, but only more probable, that Jill yesterday caught her train than that she missed it) he will not be required to say that he made an incorrect estimation if it turns out that he becomes aware (judges, perceives, understands, decides, determines, discovers, or whatever) that Jill yesterday did not catch her train. This is because what he has now committed himself to is consistent with either of the alternatives being the actual situation (and coming to be perceived by him as such). But this means that, at this moment, he doubts the truth of both propositions that represent the options, in the sense that he is not committed to the truth of either. His mind has not settled on one of them to the exclusion of the other. He has not ruled either of them out. And it is this sort of doubt that constitutes his uncertainty.

If, and as soon as, Jack comes to perceive one of the options (say "Q is the case") as being true (hence Q as actual), he commits to it (affirms it). Indeed, his coming to perceive it as true just is his becoming

committed to it (affirming it), and this process results in (indeed, constitutes) logical exclusion of the other option ("~Q is the case") from being able to inhabit Jack's representation of the actual world for as long as Jack continues to affirm that Q. If the subjective uncertainty that defines being in a state of hope resides in the object of hope being subjectively future, then this uncertainty consists of the hoping subject not being unequivocally committed to either of the two options ("Q is the case" or "~Q is the case") (i.e., by not making a positive commitment to one option to the exclusion of the other). Once the hoping subject comes to affirm one of the options (it matters not, how, or why) that equivocation or ambivalence ceases.

Affirmation of Q (or ~ Q), therefore, is what happens when Q ceases to be subjectively future for A, and A ceases to hope that Q (A's hope is fulfilled or dashed). Furthermore, this is surely what believing that consists of. Strangely enough, Day, at one point, seems to agree, for, considering how a mother might come to believe that her son has been killed in battle by being shown his body and so ceases to hope that he is still alive (!), he says that "seeing is believing: *i.e.* 'perceiving is … the acquisition of beliefs'" [8, p.90].[13] Indeed it is, especially if we understand that "seeing" or "perceiving" encompasses "comprehending that" or "realising that" as well as "sense-perception of". It is worthwhile outlining the elements of this view of belief as affirmation systematically.

Belief as Affirmation

"A affirms that Q" ↔ "A accepts 'Q is the case' is a true assertion". "A asserts that Q" ↔ "A puts forward 'Q is the case' for consideration as a true statement". However, "A asserts that Q" does not → "A accepts 'Q is the case' is a true statement". A might be guessing an answer to a question in a quiz with no idea as to whether the answer is correct. A might place a bet on Sea Fever to win the Derby without accepting that "Sea Fever will win the Derby" is a true statement. A might tell B that Q is the case just to get a rise out of B, whom A knows has a bee in her bonnet about people who assert that Q, without A accepting that "Q is the case" is a true statement. Indeed, when A asserts that Q, A might even accept that "Q is the case" is a false utterance. A could be lying.

[13] Quoting J. Heil [13, p. 238].

"A accepts that Q" ↔ "A deploys, or is disposed to deploy 'Q is the case' as a premise in A's thinking or acting for one or more purposes." So, "A accepts that Q" does not → "A accepts 'Q is the case' is a true statement". This is because A might accept "Q is the case" as a mere hypothesis or simply for the sake of argument.

When A accepts that "Q is the case" is a true statement, A treats the propositional content of "Q is the case" as part of A's mental representation of the actual world. In so doing, A treats "Q is the case" as a premise with which all other premises that A may allow to be part of A's representation of the actual world must be consistent for as long as "Q is the case" remains part of A's picture of the actual world. By affirming Q, A treats "Q is the case" as a given premise for what is the case (as against hypothetically, or for the sake of argument, or for what might be, or what probably is). In short, viewed as affirmation, it is a necessary condition of "A believes that Q" that A is committed to using "Q is the case" as an unquestioned or undoubted premise in A's current thought about what is. But this is not to say that A treats this premise as unquestionable or indubitable, nor is it to say that A regards A's commitment to it as neither revisable nor retractable.[14] Viewed functionally, to affirm (believe that Q) is to treat "Q is the case" as a true proposition, which is to act on the assumption that "Q is the case" is true (which is to prohibit action on "~Q is the case" whenever to do so is inconsistent with acting on the assumption that "Q is the case").[15] Truth can be context-dependent. For example, it can be true that Donald owns Trafalgar Square in the game of Monopoly he is playing, though he does not own Trafalgar Square in the

[14] This account entails that only beings capable of reasoning can have beliefs. A being incapable of doubting the truth of a proposition cannot have a belief, and this ability entails a capacity to understand the concept of logical contradiction. This, of course, does not entail that beliefs cannot be (subjectively) irrational. This account also requires beings with beliefs to be able to conceive of the future.

[15] This proposition is important when considering the conditions for rational belief. It entails that it is rational to believe that Q if and only if it is rational to act on the assumption that "Q is the case" is true. When the rationality of belief is at issue, it also entails that believing that Q is to be regarded as an action. Thus, it is coherent to prescribe that A believe that Q whenever it is possible for A to act on the assumption that "Q is the case" is true. It does not, however, entail that when A believes that Q, A's belief - characterising commitment to "Q is the case" (constituted by the fact that A acts as if "Q is the case" is true) - results from a choice to make this commitment, let alone a rational choice. Correlatively, that A might not be able (willing) to believe that Q when A holds values or emotional commitments that conflict with A acting on the assumption that "Q is the case" is true, does not necessarily render it impermissible to prescribe that A ought to believe that Q.

"real" world.[16] But the logic of "truth" and of "affirmation" is the same whether we are referring to the real world or a fictional world.

Concepts of Subjective Uncertainty

There are a number of different senses in which A might be certain/uncertain that Q that can be used to interpret "Asu(Q) = >0<1", the cognitive condition for "A hopes that Q". Belief as affirmation provides an interpretation of them.

1) "A does not doubt that Q". A's reasoning (theoretical or practical) is not qualified by active attention to the possibility that ~Q might be true instead. A is "positive" that Q, and does not waiver between being "positive" and "negative" that Q. For A to be certain that Q in this sense is for A to affirm Q, to believe that Q.
2) A might doubt both that Q and that ~Q, but have less doubt that Q than that ~Q. This is best described as "A is more inclined to believe that Q than that ~Q". It is fine to say that A believes that it is more probable that Q than that ~Q, but only on the understanding that this does not commit A to the idea that "Q is the case" is true. Values on the scale of subjective believability (inclination to belief), my "Asb(Q)" scale, either answer the question "Does A believe that Q or that ~Q?" or the question "To what extent is A inclined to believe that Q?"

[16] A thinking subject's ideas of the "real" world and "fictional or imaginary worlds" are both products of the subject's imagination. From the subject's point of view, the distinction between the "actual world" and "possible worlds" can be drawn both within the "real world" ("Did Dr Shipman actually murder over a thousand patients or is this only a possibility?) and within "a fictional world" ("In a game of Cluedo, did Reverend Green actually murder Mrs White, or is this only a possibility?"). The thinking subject distinguishes the "real world" from "a fictional world" in that the "real" world is that intuited by the subject as a world existing independently of the subject's senses (the phenomenology of this world being that it is posited by the subject's sense of being unable to control what the subject will experience). The picture is multi-layered, because, of course, a fictional character for A can be depicted as distinguishing between the real world and fictional worlds etc. Something can be true or believed to be true in a fictional context that is not true in a non-fictional (real world) context. In my opinion, however, necessary truths are true in all possible contexts, and include the necessary presuppositions of being able, intelligibly, to distinguish different contexts.

3) "A feels unable to doubt that Q" (which → "A does not doubt that Q" → "A affirms that Q"). "A is certain that Q" *in this sense* is a statement about the strength of A's emotional attachment to A's affirmation of Q. "A is totally attached to (A's belief that) Q" ≡ "Asa(Q) = 1". "Asa(Q) = 0" means "A is totally attached to disbelieving that Q". "Asb(Q) = 1" → "Asa(Q) = >0≤1". Statements employing the Asa(Q) scale (for all values >0) answer the question "How firmly does A believe that Q?" Asa(Q) values relate to the likelihood of A moving from affirming that Q to not affirming that Q. "A considers it to be indubitable that Q" means "A considers it categorically ought to be affirmed that Q". Unlike (1) and (2), this does not → "A affirms that Q". It is possible for A to hold completely irrational beliefs on A's own criteria. Two scales can, therefore, be generated using this sense of "A is certain that Q". Where A is presumed to affirm that Q, we have a "subjective confidence scale" (Asc(Q)). Where A is not presumed to affirm that Q, we have a "subjective justifiability" scale (Asj(Q)).

"Asc(Q) = 1" means "A believes that Q and considers this is justified beyond any possible doubt". "Asc(Q) = 0" means "A believes that Q; but A considers this is completely beyond justification or that this is completely unjustifiable".[17] "Asc(Q) = >0<1" means "A believes that Q and considers that this is justified to some degree". The Asc(Q) scale answers the question "How justified is A *in believing* that Q?"

"Asj(Q) = 1" means "A believes that Q categorically ought to be affirmed". "Asj(Q) = 0" means "A believes that Q categorically ought to be denied". "Asj(Q) = >0<1" means "A believes that neither Q nor ~Q ought to be rejected categorically". "Asj(Q) = >1/2≤1" means "A believes that there is better justification for believing that Q than for believing that ~Q". The Asj(Q) scale answers the question "How much justification does A think there is for believing that Q (or ~Q)?" Day's Asp(Q) scale is primarily an Asj(Q) scale with elements of some of the other scales. Its relations to them will be clarified later.

[17] To differentiate the disjuncts requires reference to the Asj(Q) scale. "A believes that Q; but A considers that this belief is completely beyond justification" ≡ "Asc(Q) = 0 and Asj(Q) does not apply"; "A believes that Q; but considers this belief is completely unjustifiable" ≡ "Asc(Q) = 0 and Asj(Q) = 0".

An Objection

My claim at this point is that "A hopes that Q" → "Asf(Q)" ↔ "A neither believes nor disbelieves that Q", where "A believes that Q" means "A affirms Q". I do not actually need to establish that belief is affirmation, though I will provide independent reasons for doing so later. It is sufficient for my analysis of hope that, with belief understood as affirmation, "Asb(Q) ≠ 0 ≠ 1" is necessary for "A hopes that Q". This requires, as I have shown, that "Asb(Q) = 0 or = 1" is sufficient to negate "A hopes that Q", whatever else might be formally necessary.

The obvious objection to this claim is as follows. Mere belief (as affirmation) that Q (or that ~Q) does not negate hoping that Q (hence non belief is not necessary for hope), because it is possible for A to believe that Q and still consider that it is possible that ~Q. Unless A believes that ~Q is impossible (Asπ(Q) = 1), or is certain that Q (Asc(Q) = 1), or feels that believing that ~Q is beyond A (Asa(Q) = 1), A will not have excluded ~Q from A's representation of the world, even though A believes that Q (Asb(Q) = 1). Therefore, Q will still be subjectively future for A!

This objection relies on an equivocation in the statement "It is possible for A to believe that Q and still consider that it is possible that ~Q". The statement is true if, in the sub-proposition "It is possible for A to believe that Q", "that Q" is part of A's representation of the actual world, while, in the sub-proposition "It is possible for A to consider that ~Q", "that ~Q" is part of A's representation of a possible world.

The statement is false if "that Q" and "that ~Q" are both held to be part of A's representation of the actual world. This is because, while A's belief that Q does not exclude ~Q from A's representation of a possible world (because "Asb(Q) = 1" does not → "Asπ(Q) = 1", or "Asc(Q) = 1", or "Asa(Q) = 1"), it does exclude ~Q from A's representation of the actual world (for "Asb(Q) = 1" → "Asb(~Q) = 0"). A, in believing that Q is actual now, does not also believe that ~Q is possible now (at the same time as believing that Q is actual), and so does not hope that Q now, because this requires A to entertain both that Q and that ~Q now.

That "Asb(Q) = 1" does not → "Asπ(Q) = 1", or "Asc(Q) = 1", or "Asa(Q) = 1" does not show that "A hopes that Q" does not → "Asb(Q) ≠ 0 and≠ 1". What "Asπ(Q) ≠ 1", or "Asc(Q) ≠ 1", or "Asa(Q) ≠ 1" signify when "Asb(Q) = 1", is that A envisages the possibility that A could change A's mind that sb(Q) =1.

While it is true that focussing on this possibility might lead A to cease to believe that Q (which will be subject to influence by how much A desires that Q, by how much below 1 Asc(Q) ≠ 1 or "Asa(Q) ≠ 1 is, and by how rationally motivated A's beliefs are) this is not necessarily the case. In any event, the point is that only if and when A's recognition that π(Q) ≠ 1, or c(Q) ≠ 1, or a(Q) ≠ 1 leads to "Asb(Q) ≠ 0 and ≠ 1", will A hope that Q. Failing this, "Asπ(Q) ≠ 1", or "Asc(Q) ≠ 1", or "Asa(Q) ≠ 1" can (at most) only place A in a state of hope that A will cease to believe that Q, equivalent to A hoping to be able to hope that Q.

My reply, then, is that unless they lead to A ceasing to believe that Q, states of mind like "Asπ(Q) ≠ 1", or "Asc(Q) ≠ 1", or "Asa(Q) ≠ 1" do not signify that A continues to hope that Q.

From the point of view of the hoping subject, A, they only serve to generate a secondary object of hope, "A hopes that A can come to hope that Q". From a third person perspective "Asπ(Q) ≠ 1" is cognitively necessary *and sufficient* only to see Q as a possible (intelligible) object of hope for A, not as an actual one.

This distinction between primary and secondary hopes is not merely an ad hoc device to rescue my analysis. It is something that the objection itself must make unless it is to lead to an infinite regress. This being the case, my thesis that "A hopes that Q" → "Asb(Q) ≠ 0 and≠ 1" is not just a consistent view of "Q is subjectively future for A", but a necessary one.

According to the objection, "Asb(Q) = 1" is not sufficient to negate "A hopes that Q" because, unless a stronger modal condition is operating, A has not altogether excluded the possibility of not believing that Q.

If, contrary to my claim, we suppose that this refutes my thesis, then we must note that all the alternative hope negating conditions involve believing (estimating, considering, judging, thinking, perceiving) that Q is impossible, or conclusively confirmed, etc.

As such, if it is claimed, e.g., that "Asπ(Q) = 1" negates "A hopes that Q" then this claim must be false. This is because "Asπ(Q) = 1" merely says that A believes that Q is certain (~Q impossible). Unless A believes that A's belief that Q is certain, is itself not capable of being false, A has not excluded the possibility that ~Q is the case, and so on ad infinitum. The consequence is that "A hopes that Q" can never be negated except by "A does not desire Q". But this regress can only be stopped by distinguishing primary and secondary hopes and requiring "Asb(Q) = 0 (or = 1)" to be sufficient to negate "A hopes that Q".

"Asπ(Q) ≠ 0 and≠ 1" Is Redundant

"Asπ(Q) = 0 (or = 1)" negates "A hopes that Q". However, I have now established that "Asb(Q) = 0 (or = 1)" is also sufficient to negate "A hopes that Q". This entails that both "Asπ(Q) ≠ 0 and ≠ 1" and "Asb(Q) ≠ 0 and ≠ 1" are both necessary for "A hopes that Q". This, however, can only be the case if the two conditions mutually entail each other. This is because unless necessary conditions mutually entail each other (and are hence just one condition) they cannot be self-sufficient (only jointly sufficient).

It is clear that "Asπ(Q) = 0 (or = 1)" → "Asb(Q) = 0 (or = 1)". A cannot believe that Q is impossible (or certain) and not believe that ~Q (or that Q). Therefore, "Asb(Q) ≠ 0 and ≠ 1" ≡ "~[Asb(Q) = 0 (or = 1)]" → "~[Asπ(Q) = 0 (or = 1)]". On the other, hand, "Asπ(Q) ≠ 0 and ≠ 1" does not → "Asb(Q) ≠ 0 and ≠ 1", because A can believe that Q when A believes that Q is neither certain nor impossible. Since, provided that A has Q in mind, "~[Asπ(Q) = 0 (or = 1)]" → "Asπ(Q) ≠ 0 and≠ 1", the latter condition is then cognitively redundant as a necessary condition for "A *does* hope that Q". While it is formally necessary for "A hopes that Q", it is not sufficient, and it is satisfied whenever "Asb(Q) ≠ 0 and ≠ 1" is satisfied. As I explained in the previous sub-section, however, "Asπ(Q) ≠ 0 and ≠ 1" is a necessary condition for "A could hope that Q" (when A does not hope that Q). There is, however, no harm in specifying the redundant condition, which, in effect, specifies the "in mind" qualification.

Problems with Day's View of Belief and the Justification of Belief as Affirmation

To remind ourselves, according to Day, "A hopes that Q" → "Asp(Q) = >0<1" meaning "A considers that Q and ~Q both have limited probability". "Asp(Q) = 0" ≡ "A is certain that ~Q". "Asp(Q) = 1" means "A is certain that Q" [8, p.32] ≡ "A is confident that Q" (see ibid., but contrast [**8**, p.65]) ≡ "A is convinced that Q" ≡ " A knows that Q" [8, p.72]. "Asb(Q) =>0>1/2" means "A suspects that Q" [8, p.72]; "Asp(Q) = >1/2<1" means "A believes that Q" [8, p.72].

This is counterintuitive. For example, if "A suspects that Q" means "Asb(Q) =>0>1/2" and "A believes that Q" means "Asb(Q) =>0>1/2", then

"A suspects that Q" means "A believes that ~Q". If "A believes that Q" means "Asb(Q) =>0>1/2" then, because "A knows that Q" means "Asp(Q) = 1", "A knows that Q" → "A does not believe that Q", which also entails that knowledge cannot be justified as true belief. And none of this is improved if we replace "equivalence" with "→".

If belief is affirmation, then the basic problem with Day's view is that A must either believe that Q is probably the case or be certain that Q is the case. There is no space for something in between. "A perceives (comprehends) that Q is actually the case" is the central idea when belief is affirmation. But "Q is actually the case" either has no place at all in Day's scheme or else must be reduced to "Q is probably the case" or to "(A's) certainty that Q is the case". "A believes that Q" ↔ "Asp(Q) = >1/2<1" requires us to accept that "A believes that Q" means "A considers that the proposition 'Q is the case' is more likely to be true than not; but A is not certain that it is true". If we accept this, then we must give up the idea that "A believes that Q" → "A accepts that the proposition 'Q is the case' is true". But to believe that something is the case is distinct from not being certain that it is the case but believing that it is probably the case. If A believes (on 20 November) that Q is actually the case (regardless of whether or not A is certain that Q is the case), and on 21 November comes to believe that Q is not the case, A must declare that A held a false belief on 20 November. On the other hand, if A (on 20 November) merely believes that Q is probably the case, and on 21 November comes to believe that Q is not the case, A has no need to accept that A held a false belief on 20 November. What A affirmed on 20 November was that Q is probably the case. In so doing, A neither affirmed that Q is the case nor that ~Q is the case. So, by affirming (on 21 November) that ~Q is the case, A does not deny what A affirmed on 20 November.

To this, of course, it will be responded that that I have not actually justified the thesis that belief is affirmation, and the idea that knowledge is justified true belief is not sacrosanct, etc. True, but there are problems with Day's view which do not rest on assuming that belief is affirmation, which can only be remedied by holding that belief is affirmation.

The Need for Belief as Affirmation

Day acknowledges that if "A believes that Q" entails "Asp(Q) = >1/2<1", then belief that Q "entails the existence of a second-order-belief about the

probability of Q“ [8, p.73], namely, the belief that the probability of Q = >1/2<1. But it follows that “A believes that the probability of Q = >1/2<1” → “A believes that the probability is >1/2<1 that the probability of Q= >1/2<1“. This sets up an infinite regress. Day denies this, by rejecting the claim that the existence of the second order belief he identifies “entails the existence of a third-order belief about the probability of the belief about the probability of Q” [8, p.73]. He appeals to an analogy with a desire not to smoke tobacco, claiming that, while there can be a desire not to desire to smoke tobacco (which is true), it does not follow that there must be “a third-order desire to desire to desire not to smoke tobacco” [ibid] (which is also true). But the analogy is inappropriate, and this is because, while there can be a desire to desire not to smoke tobacco, it is not true (and Day does not claim that it is true) that a desire not to smoke tobacco entails a second order desire to desire not to smoke tobacco. The claim of entailment is what sets up the regress. Indeed, as I have noted,[18] he is actually committed to the idea that “Asp(Q) = >1/2<1” is not only entailed by “A believes that Q”, it specifies the meaning of “A believes that Q”.

It might, perhaps, be thought that this problem can be avoided by specifying that “A believes that Q” → “Asp(Q) = >1/2≤1”, which would avoid some of the counter-intuitive aspects of Day’s view. However, this still produces numerous problems. If A believes that p(Q) = 1, then he can only do so if he is certain that p(Q) = 1. Otherwise he believes that the probability of p(Q) = 1 is >1/2<1, which means that it is >1/2<1. So, he can only ever be certain that Q, when he is certain that he is certain that Q. This does not set up a regress. It means that he can only believe that Q when he is certain that Q.

But if this move is not made, then how is it ever possible for A to be certain that Q? To be certain that Q, A must estimate p(Q) = 1. But to estimate, judge, perceive etc. that Q is to believe that Q. So, “Asp(Q) = 1” → “Asp(Q) = >1/2<1”, which is self-contradictory. There is no need to go on pointing out the paradoxes that are generated by Day’s view: the problems arise from trying to find the essence of having a belief in attribution of a degree of confirmatory certainty or uncertainty. This, in effect, amounts to the reduction of to the idea of “the actual” to “the certain” or “the probable”.[19] The absurdities that result show that any adequate theory of belief must be able to distinguish (a) “A

[18] See n.11 supra.

[19] This is an example of what Roy Bhaskar has called “the epistemic fallacy”, the “idea that being can always be analysed in terms of our knowledge of being” [7, p. 36]. The Verification Principle of Logical Positivism also involves this fallacy.

perceives that Q is actually the case" from (b) "A perceives that Q is probably but not actually the case" from (c) "A is certain that Q is actually the case". This requires the distinctions between different ideas of certainty and possibility constituted by my Asb(Q), Asc(Q), Asj(Q), and Asπ(Q) modalities, as co-ordinated by the idea that belief is affirmation.

Day's Justification for His View of Belief Is Inadequate

The only direct argument Day offers in support of his analysis of "A believes that Q" is contained in the following passage.

> According to what may be called the classical theory of Belief, as propounded by Locke and Price, Belief is a genus, the different species of which are its degrees, which range from Suspicion at the bottom of the scale to Conviction at the top of it. But this seems to me to misrepresent the way in which the verb "believe" works. Thus it is correct to say "The police suspect that Sykes did it, but they do not yet believe it". (They need more and better evidence for that). Again, if John is convinced that the Earth is flat, he will not say that he believes this; he will claim to know it. … Belief does not comprehend Suspicion, Conviction *etc.*; it is just one propositional attitude among many, just as they are. The differences between e.g. Suspicion, Conviction and Belief are as follows: (i) "A suspects that Q" entails "sp(Q)>0<1/2"; (ii) "A is convinced that Q" entails "sp(Q) = 1"; and (iii) "A believes that Q" entails "sp(Q)>1/2<1". One cannot correctly say "A believes that Q although he believes that Q is unlikely (*i.e.* that Q has probability >0<1/2)" [8, p.72].

I do not subscribe to what Day calls the "classical theory" of belief. In my view, "A believes that Q" means "A is unequivocally committed to the truth of 'Q is the case'"; "A suspects that Q" means "A leans towards unequivocal commitment to the truth of 'Q is the case', but A does not yet have such a commitment"; while "A is convinced that Q" (meaning by this what Day does, that A is certain that Q or knows that Q, but noting that in ordinary usage this can also mean simply that A believes that Q)[20] means "A believes that Q and believes that this cannot possibly be doubted". This attributes three different

[20] Day has it that "A has the conviction that Q" ↔ "Asp(Q) = 1". I find it more natural to say that "A has the conviction that Q" states that "A is positive (not + or -) that Q" rather than "A is normatively certain that Q" and merely states that "A believes that Q"; but this is incidental to the matter at hand.

modalities to the three attitudes. They are not, as Day claims the "classical theory" holds, different degrees on a scale with a single modality. This is ironic, because it is Day's own explication of what these three attitudes involve that puts them all on the same scale, involving just one modality, that of subjective evidential probability.

It is true that Day does not hold that suspicion and conviction are degrees of belief: he holds, instead, that suspicion, belief and conviction are degrees of subjective probability.[21] So, what is his justification for this? The idea that when the police merely suspect Sykes did it, they do not yet believe it, and the claim that if John is convinced that the Earth is flat he will not claim to believe it, but claim to know it. Well, the first claim is true. "A believes that Q" →"A does not (merely) suspect that Q". And this does show that suspicion is not a kind of belief but a state leading up to belief. As I have said (see my Asb(Q) and Asj(Q) scales) this can be scaled. No disagreement here, except that for Day it is necessarily a justificatory scale.

The second claim, however, is different. John might not claim to believe that the Earth is flat, if he is certain that it is flat, but instead claim to know that it is flat. But this only shows that "A knows that Q" → "A does not believe that Q" if it is necessarily true that if John said "I believe that the Earth is flat — in fact, I know that it is flat" he would be contradicting himself or changing his mind very quickly. But this is not the case if a claim to knowledge is (or John thinks it is) a qualification of a belief (e.g., the qualification that the belief is a justified true belief). Nothing in this example precludes interpreting John's statement in this way, unless it is presumed (when this is just what the example is supposed to be demonstrating) that "Belief does not comprehend … Conviction [i.e., being convinced]" in a way that does not permit an sp(Q) = 1 to be associated with a belief.

So, what do I make of the paradox alleged in the final sentence of the quoted passage? Certainly, *if* "A believes that Q" ↔ "Asp(Q) = >1/2<1", then to hold "A believes that Q, but Asp(Q) = >0<1/2" is self-contradictory, because this asserts "Asp(Q) = >1/2 and <1/2". But Day has given no justification for "'A believes that Q' ↔'Asp(Q) = >1/2<1'" apart from a question-begging interpretation of his example about John being convinced that the Earth is flat.

[21] Day's claim that "Belief does not comprehend Suspicion", which is correct, simply does not square with his view, which, as I have already pointed out entails "A suspects that Q" → "A believes that ~Q".

So, how does the idea that A believes that Q but thinks that Q is unlikely fare in my theory? In my scheme, "A thinks that Q is unlikely" is ambiguous as between "Asb(Q) = >0<1/2" and Asj(Q) = >0<1/2". On the one hand, to say that A believes that Q (Asb(Q) = 1) but considers Q to be unlikely (less believable than ~Q) (Asb(Q) = <1/2) is contradictory. But this provides no support for Day's view, because it is also contradictory to say that Asb(Q) = 1 (to say that that "Q is the case" is true) and that Asb(Q) = >1/2 <1 (to say that "Q is the case" is only probably true). On the other hand, to say that A believes that Q, yet considers that Q is unlikely (less justified than ~Q), is not self-contradictory. To say that A believes that Q is to say that Asb(Q) = 1, and to say that A considers Q to be less likely (less justified) than ~Q is to say that Asj(Q) = >0<1/2 . This is only a contradiction if it is a contradiction to say "A believes that Q but considers that A ought not to believe that Q". This can only be a contradiction if it is a contradiction in terms for A ever to consider that A might hold what A considers to be a rationally unjustified belief or even one lacking a rational justification. This is simply not so, and Day, himself, points out that "It is no news that human beings are not always logical" [8, p.60]. Quite! Day's analysis of belief is wholly dependent on over-rationalising the process and state of having a belief.

ANSWERING DAY'S CHALLENGE

This brings us, finally, to Day's direct argument that an analysis of "hope" requires reference to subjective probability. In effect, he agrees that "A hopes that Q" ↔ "A desires that Q; and Asu(Q) = >0<1", and the reason he gives why "Asu(Q) = >0<1" must be interpreted as "Asp(Q) =>0<1" is that the analysis of, e.g. "A has high hope that Q", requires reference to subjective probability.

It is true that "A has high hope that Q" requires reference to subjective probability. But it does not follow that "A hopes that Q" ↔ "A desires that Q; and Asp(Q) =>0<1". This is because I (and proponents of the orthodox view) can claim that the correct analysis of "A has high hope that Q" is "A hopes that Q; and A estimates the probability of Q as high". Here, "A hopes that Q" is not to be analysed as "A desires that Q; and Asp(Q) = >0<1", but (in my view) as "A desires that Q; and Asb(Q) ≠ 0 and ≠ 1", and in the orthodox view as "A desires that Q; and Asπ(Q) ≠ 0 and ≠ 1". In my view, subjective probability seen as compatible with hope is measured in terms of "Asb(Q) =

>0<1", while orthodox views are likely to measure it in terms of "Asj(Q) = >0<1" — rather than Day's "Asp(Q) = >0<1" — as they will not adopt Day's view of belief.

Hence, in my view, "A has high hope that Q" ↔ "A desires that Q; Asb(Q) ≠ 0 and ≠ 1; and Asb(Q) is high". The negative point is that, while there can be no doubt that "A is hopeful that Q" cannot be analysed without reference to (subjective) probability, it simply does not follow that "A hopes that Q" cannot be analysed without reference to probability. More is required to show that.

The positive point is that the essence of hope lies in the intersection between desire and non belief, or (in the orthodox view) that between desire and subjective possibility, and that (in either case) subjective probability bears only on the degree of expectation[22] that A has that Q will be realised. In other words, subjective probability is the cognitive aspect of degree of expectation, not the cognitive aspect of hoping or even part of the cognitive aspect of hoping.

[22] If A believes that Q then A expects Q. Since belief negates hope, A cannot hope that Q and expect Q at the same time. To expect Q is to depict Q as what will be, not as what might or might not be. Like believing that Q, there are no degrees of expecting Q. However, A can have degrees of expectation that Q, which track how probable A thinks Q is, parallel to the degrees to which A suspects that Q. There is an ambivalence here, similar that affecting "A is confident that Q" and "A is certain that Q". In some contexts, we think that if A has any doubts that Q then A is not confident that Q. Being confident does not admit of degrees. But we do also speak of A being confident (or having confidence) to varying degrees. I suggest that the latter is actually a way of saying that A is lacking in confidence to varying degrees, enabling us to say how far from being confident in the all or nothing sense A is. (Day himself recognises this ambivalence without seeming to be aware of it, when he says that "A is confident that Q" → "Asp(Q) = 1" [8, p.32], but "Asp(Q) = >1/2<1" [8, p.65]). Much the same can be said about "certainty". We sometimes say that A is either certain or A is not. But we do also speak of A being certain to varying degrees. But is it not really that it is A's uncertainty, A's falling short of being certain, that varies in degrees? Whatever one makes of this, it is important to recognise these subtleties. Otherwise, they become a source of misunderstanding and equivocation.

In response to the charge that the Christian idea (as, e.g., expressed in the Anglican Burial Service) that Christians have a sure and certain hope and expectation of the life to come is, if meant literally, a self-contradiction, A. Phillips-Griffiths [24], claims that hope is "certainty about the hypothetical". Christians are certain that God will save them *if* they carry out their Christian duties. The only uncertainty is about whether or not they will carry out their duties, not about what God will do. But, while Phillips-Griffiths renders the words of the Burial Service coherent, this does not answer the charge at all. His analysis has Christians expect that God will save them if they carry out their Christian duties and hope that they will carry out these duties. It does not portray Christians as hoping that God will save them if they carry out their duties. Contradiction is avoided by giving the hope and the expectation different objects.

The cognitive aspect of hoping is non belief (my view) or subjective possibility (the orthodox view). But neither non belief nor subjective possibility can vary in degree. Something is possible or it is not. A believes that Q or A does not believe that Q. While there can be degrees of leaning towards believing that Q (or towards disbelieving that Q) there cannot be degrees of believing that Q (or of non believing or disbelieving that Q).[23]

So, while I agree with Day that hope can vary according to degrees of the conative aspect, I do not agree that it varies according to degrees of the cognitive aspect, and this is simply because the cognitive conditions for hope as such do not vary in degrees.[24]

IMPLICATIONS FOR KANT'S MORAL ARGUMENT FOR THE EXISTENCE OF GOD

Immanuel Kant famously argued that even though we cannot know whether or not God (conceived to be omnipotent and perfectly good) exists [18, A742 B 770; A 592-630 B620-658] God's existence is "postulated" by the moral law. Because the moral law is "connected (completely a priori*)* with the concept of the will of a rational being as such" [19, 4: 426] (i.e., morality is a requirement of pure practical reason), belief in God is rationally necessary in the strictest sense. Rational beings with a will ("agents"), i.e., those who pursue ends as reasons for their actions, contradict that they are agents if they do not consider themselves bound by the moral law [19, 4: 428-429]. Consequently, they must believe that God exists, not only to be consistent with any commitment they have to the moral law, but in order to be consistent with the idea that they are agents. Therefore, while theoretical reason requires agnosticism, pure practical reason requires theism.

[23] Of course if "Q is the case" is a compound proposition that can be broken into a number of discrete propositions, then it is possible to believe that Q to a degree if what is meant is that some of the component propositions are believed whereas others are not. But this is trivial.

[24] In Day's view, when A desires Q, A's hope for Q increases as Asp(Q) increases within the interval 1/ 2 to 1. In my view, A's degree of hope remains constant and, *if anything,* A's attitude towards Q becomes *less* one of hope and more one of expectation (i.e. A leans more away from hope towards belief as the subjective probability of Q for A (short of amounting to belief) increases..

I understand Kant's argument, at least as presented in *Critique of Practical Reason* [20, 5:122-126], to be as follows:

1) If the moral law were fully complied with and never violated, happiness and worthiness for it would be in complete harmony. Such a state-of-affairs is the *"summum bonum",* the highest good.
2) The moral law "postulates" the *summum bonum*: i.e., under the moral law, the *summum bonum* is the "final" end of all action, which, ideally, ought to exist.
3) The moral law requires agents not only to want the *summum bonum* to be realized; it requires them to do whatever they can to bring it about. The *summum bonum* is a necessary object of the will.
4) Unless God exists (and agents are immortal),[25] the *summum bonum* is unrealisable.[26]
5) Since "'ought' implies 'can'", agents may take the moral law to prescribe that they pursue the *summum bonum* only if they assume that God exists.

Therefore

6) Agents who regard themselves as bound by the moral law ought, in consistency with this commitment, to believe that God exists.

Combined with Kant's view that commitment to the moral law is a requirement of pure practical reason, this result is sufficient to ground practical theism, the thesis that it is rationally necessary in the strictest sense for agents to believe that God exists.

However, Kant does not think that this proves that God exists [20, 5: 138]. Practical reason requires agents to have "faith" or "rational belief" that God exists; but they do not, thereby, know that God exists [20, 5: 144-146; 18, A829 B857]. In *Critique of Pure Reason*, he states that God's existence is certain, but this certainty is moral certainty not logical certainty [18, A829 B857]. When he says that belief in God's existence is certain, he means that it is necessary for agents, qua thinking of themselves as agents, to believe that God exists. However, since the requirement to believe that God exists is driven by the moral law (as a requirement of pure practical reason), he must also

[25] To simplify presentation, I will not repeat the immortality condition, but take it as read.

[26] I think this claim is correct, but I will not attempt to defend it here.

claim that agents morally ought to believe that God exists (i.e., morally ought to treat "God exists" *as* true, which is to treat it as a premise for their thought and action), which makes it wholly unsurprising that in *The Metaphysics of Morals*, he declares that "to have religion is a duty of man to himself" [21, p.238]. In effect, practical reason via the moral law generates a maxim, "I *will* that there be a God!" [20, 5: 143], which is to say, "Act as if there were a God!" meaning "Act on the presumption that the *summum bonum* is, cosmologically, the purpose of existence!"

A Standard Objection to Kant's Argument

A standard objection is that (3) is false because "The *summum bonum* ought to be!" is not *a* command for action, but an "'ought' of evaluation": eventuation of the *summum bonum* is good for finite agents, but not a duty of finite agents because it is not within their power (individually or collectively) to bring it about.

The moral law only commands that finite agents act in accordance with the moral law, which they can do, whether or not God exists. In the words of Lewis White-Beck, the moral law "as an imperative … is a command only that we seek virtue, let the eschatological chips fall as they may" [3, p.275].

Consequently, (3) must be replaced with something like

> 3) Under the moral law, agents must want the *summum bonum* to be realized *and* do nothing contrary to its realization, for what they ought to desire (would desire if they were fully rational) and the ends they ought to pursue must be in harmony. *In this sense only* is the *summum bonum* a necessary object of the will.

In this sense, God is also a necessary object of the will; but if only in this sense, this means no more than that, under the moral law, agents must want God to exist.

With the moral law being rationally necessary, it follows only that it is rationally necessary for agents to want God to exist. Of course, if the world, in the cosmological order of things, is ordered as pure practical reason dictates it ought to be, then God necessarily exists. However, only if reason requires agents to think that the world is necessarily ordered as it ought to be, does it require agents to believe that God exists.

But, unless agents know that God exists (which they cannot), they have no good reason to suppose that the world is necessarily ordered as it ought to be. There is a circularity here that cannot be broken.

Kant's error, on this account, is that he equivocated between the *summum bonum* as an object of rationally required desire and the *summum bonum* as a morally required goal for action.

Nevertheless Kant's Argument Shows that Agents may Not be Atheists

Even if this is so, it is a mistake to conclude that Kant's considerations are neutral as to what rational agents may believe about God. Atheism requires agents to characterize the moral law (and, indeed practical reason) as requiring them to want something to exist that cannot possibly exist. This is because the moral law requires them to want the *summum bonum* to be brought about, and given the realization that God must exist if the *summum bonum* can possibly be brought about, to believe that God does not exist is to believe that the *summum bonum* cannot possibly be brought about.

Now, if "'ought' implies can" applies to "'oughts' of evaluation" (as well as to "action-directing 'oughts'"), meaning that it is irrational to judge that something ought to exist if one supposes that it is impossible for it to exist, then the untenability of atheism on moral grounds is clear. The moral law requires agents to judge that the *summum bonum* ought to be, so agents cannot (in consistency with the idea that they are bound by the moral law) suppose that the condition required for it to be, God's existence, is not in place. Indeed, on the basis that the moral law is dialectically necessary, agents may not believe that God does not exist for any reason, because there are no rational grounds for believing that God does not exist more rationally compelling than those requiring agents to respect the moral law.[27]

However, rather than rely directly on the claim that "'ought' implies can" does apply to "'oughts' of evaluation", I will offer two other arguments against atheism. The first argument is that atheism undermines respect for the moral law and practical reason by challenging the idea that the moral law and pure practical reason are categorically binding.

[27] The strongest arguments for atheism allege that an omnipotent perfectly good God cannot tolerate the existence of manifest evil in the world. This problem is tackled by theodicy, which is a large topic. I believe (and will here suppose) that the problem can be solved.

If God does not exist, and we (agents) are not immortal, then our lives and actions have, in the final scheme of things, no significance. In the words of the Anglican Burial Service, our existence is no more than a journey from "earth to earth, ashes to ashes, dust to dust!" In the words of Johannes Brahms' *German Requiem: "*Denn alles Fleisch es ist wie Gras und alle Herrlichkeit des Menschen wie des Grases Blumen" ("For all flesh is as grass, and all the glory of man as the flower of grass"). If so, then even though morality and practical reason do, on their own terms, require us to attach categorical significance to ourselves, both have no ultimate significance in themselves and it is a deceit that their unconditional requirements are to be respected categorically. Indeed, Kant presses this very argument when he maintains that

> righteous man (like Spinoza) who takes himself to be firmly convinced that there is no God and … no future life … [must, in the final analysis, view himself not as an end-in-itself, but as destined for] the abyss of the purposeless chaos of matter. … [This] weaken [s] the respect, by which the moral law immediately influences him to obedience, by the nullity of the only idealistic final end that is adequate to its high demand (which cannot occur without damage to the moral disposition) [22, 5:452].

The second argument is that atheism renders the moral law's requirements incoherent. Agents ought to be unhappy if their rationally required desires are not fulfilled. Not to get what we ought to desire is not merely just cause for dissatisfaction, but demands dissatisfaction. Under the moral law, we categorically ought to desire that not only ourselves, but all others, not be victims of violations of the moral law and we categorically ought to be unhappy when any agent is the victim of uncompensated injustice.

However, if we suppose that God does not exist, so that the *summum bonum* cannot be brought about, we must suppose that agents will inevitably suffer uncompensated injustice. We must, then, characterize the moral law and pure practical reason as unconditionally requiring us to be unhappy, whether or not we do our duty under the moral law. However, it is because the moral law postulates as an ideal good that we ought to achieve happiness if we do our duty that it postulates the *summum bonum.* Therefore, to believe that God does not exist is to portray the moral law as self-contradictory: it judges that we ought to be unhappy whether or not we do our duty, yet judges that we ought to be happy provided only that we do our duty.

But This Does Not Entail that Agents Must Believe that God Exists

This suggests, as an alternative to the standard account, that Kant's error in arguing that the moral law requires agents to believe that God exists is that he concludes from the valid inference (resting on the *summum bonum* as merely an object of rationally required desire) that agents may not believe that God does not exist that they must believe that God exists. For, while the negation of "God does not exist" is "God exists", the negation of "I believe that God does not exist" is not "I believe that God exists", but "I do not believe that God does not exist". The latter proposition is compatible with both "I believe that God exists" and "I do not believe either that God exists or that God does not exist". In short, if we may not be disbelievers (atheists), we need not be believers (theists). We may be non-believers (agnostics) instead.

I hesitate to suggest that this was Kant's error, because Kant was, at least in principle, aware of these distinctions [18, A503 B531; A791 B819]. In any case, we must at this point conclude that agents may be theists or agnostics, but not atheists.

Theism Is Also Incompatible with the Idea that Agents Are Bound by the Moral Law

However, closer examination reveals that, under the moral law, it is not permissible to be theists either. Kant insisted that the moral law is not known on the basis of religious belief. Not only was he confident that agents can be certain that they are bound by the moral law on purely *a priori* grounds, he was adamant that the only basis they have for the idea that God is omnipotent and perfectly good is the moral law [19, 4: 408-409]. For Kant, God's existence is not a transcendental condition of the possibility of morality, but an inference from the existence of morality. Therefore, anything agents say about God must be consistent with the transcendental conditions of the possibility of morality.

Now, amongst these conditions are those that are necessary for morality to be intelligible, and Kant was aware that intelligible subjects and objects of the moral law, viewed as an imperative, must perceive themselves to be vulnerable both in being able to obey/disobey the moral law [19, 4: 414] and

in being capable of being harmed morally.[28] However, *if* God exists then the *summum bonum* will necessarily be realised. As Leibniz proclaimed [23, p.27], and Voltaire lampooned in *Candide* [29], if God exists then all must be for the best and this must be the best of all possible worlds, otherwise "God" cannot be both omnipotent and perfectly good. But this implies that no sincere and sane theist who understands the concept of God given by the moral law, having in mind God the all-loving savior — who guarantees full redress, and ultimately salvation for all — could possibly think that agents need the protection of a categorical imperative. In the words of Psalm 23,

> Yea, though I walk through the valley of the shadow of death, I will fear no evil: for Thou art with me … . Surely goodness and mercy shall follow me all the days of my life: and I will dwell in the house of the Lord forever [King James Version].

The idea that our actions can make a difference to the ultimate order of things becomes vain. The bringing about of the *summum bonum* is God's responsibility, not ours. Our only responsibility is to obey the moral law. But that is not enough to bring about the *summum bonum.* Indeed, although our willing conformity is formally necessary (for the *summum bonum* will not be realized while there are transgressors), God, by definition, will bring about the *summum bonum* no matter what. In addition, all "harms" suffered must eventually be seen by their victims to be justified as being for the best in this the best of all possible worlds. And, with the *summum bonum* involving eternal salvation and redress, its achievement must constitute nothing less than the end of all harm and the end of any further need for the moral law as an imperative. In short, from the perspective of the achieved *summum bonum* there can be no moral harms at all.

Nor could a comprehending, sane, and sincere theist, having in mind God, the omnipotent and omniscient Judge, possibly be tempted to disobey the moral law, which makes a mockery of any idea of freedom. And Kant reasons in just this way when he asserts that an ability to prove that God exists would be disastrous for morality. If agents knew that God exists,

[28] Kant has surprisingly little to say about this, but recognition of it is implicit in his depiction of the 'starry heavens above' as symbolizing a material world devoid of meaning and thereby threatening to annihilate not only agents' physical selves but any pretensions to significance they might have [20, 5:161-162].

> Most actions conforming to the law would be done from fear, few would be done from hope, none from duty. The moral worth of actions … would not exist at all. The conduct of man, so long as his nature remained as it now is, would be changed into mere mechanism [20, 5:147].

In short, those who were even momentarily tempted to transgress would display a lack of reason that would excuse them from responsibility for their actions. In effect, according to Kant, the idea that God's existence is knowable conflicts with the transcendental conditions of the possibility of the moral law presenting itself as a categorical imperative. However, Kant thinks that this conflict between theism and morality applies only to the supposition that God's existence can be proven, not to practical theism.

But why? The objections to theism just cited (including Kant's own) rest on the practical effect of believing (i.e., supposing it to be true) that God exists, not on the idea that the proposition that God exists is proven to be true (hence, certainly true). And, even if it did rest on supposing it to be certain that theism is true, it would still apply to Kant's practical theism, according to which agents are morally required to be certain that God exists [18, A829 B857].

The Implication is that Agents must be "Hopeful" Agnostics

It follows that, while theoretical reason merely does not enable agents to know whether or not God exists, practical reason positively requires them not to believe either that God exists or that God does not exist, yet to want God to exist. How are we to describe such a state of mind?

On the conception of "hope" that I have argued for in this chapter, the conclusion we have reached is that agents must be agnostics in thought and action who, no more and no less, hope that God exists and fear that God does not (for hope and fear are opposite sides of the same coin).

Those who want God to exist but believe that God does not exist are not in a state of hope. If they consider that their belief that God does not exist might be mistaken, the only thing they can be properly said to hope is that they are mistaken in their belief that God does not exist, but unless (and until) this perceived possibility leads them to stop believing that God does not exist, they are (or ought to be) in a state of despair; for the world they see is very far from the best of all possible worlds. On the other hand, with parallel qualifications, those who want God to exist and believe that God does exist are in a state of

joyful expectation that all will (not might) prove to be for the best in this the best of all possible worlds.

In the Preface to the 2nd edition of *Critique of Pure Reason*, Kant declares, "I have … found it necessary to deny knowledge, in order to make room for faith." [18, Bxxx]. What he should have concluded was "I have denied knowledge and faith in order to make room for hope and fear, and, thereby, for morality!" And this is because there is nothing more poisonous to the moral disposition than loss of hope and its mirror fear, whether this be at the gain of despair (whose constituents, feelings of overwhelming impotence, insignificance and pointlessness, annihilate hope) or at the gain of joyful expectation (whose components, feelings of final security, significance and purpose, nullify fear).[29] We must, however, note that Kant (correctly) characterizes hoping in *Critique of Pure Reason* as standing "in the same relation to the practical and the law of morality as knowing and the law of nature to the theoretical knowledge of things". But he then declares that, where something is necessary for something to happen, hoping "arrives … at the conclusion that something is … because something ought to happen as against that something is … because something happens" [18, A805-806 B833-834].

But if the relation is the same in both cases, Kant should say that "something ought to be because something ought to happen". If he is saying that where X is necessary for something that ought to happen, agents ought to hope that X will happen and this means they ought to believe that X will happen, this is clearly invalid and involves an idiosyncratic idea of hope (though it looks consistent with the way in which he tries to justify the postulate that God exists). However, he might just mean that in this scenario agents ought to "have faith" that X is the case and this means they ought to hope (according to my definition) that X will happen (involving no more than judging that X ought to be).[30] If we follow through on this, then we must

[29] Though they lack a sound transcendental foundation, there are some excellent phenomenological accounts of these matters in the tradition of existential psychology. See, e.g., Rollo May [16], and Erich Fromm [10].

[30] Such a reading is not implausible if we attend to how Kant formulates his argument for God in *Critique of Pure Reason*. There, Kant asks, "If I so behave as not to be unworthy of happiness, may I *hope* [my emphasis] thereby to obtain happiness?" [18, A809 B837]. In answering this question he says,

> "The alleged necessary connection of the *hope* [my emphasis] of happiness with the necessary endeavour to render the self worthy of happiness [this endeavour being obedience to the moral law] … can be counted upon only if a Supreme Reason, that governs according to moral laws, be likewise postulated as underlying nature as its cause" [18, A810 B838].

understand Kant's practical 'belief' or 'faith', not as rationally necessary propositional belief, but as rationally required hope. Then, his claim that God's existence is morally certain must be taken to mean that hope that God exists is morally required. If so, then what Kant actually means by practical theism, is hopeful agnosticism.

CONCLUSION

I have argued that "A hopes that Q" ↔ "A desires that Q; and A neither believes nor disbelieves that Q". Unbelief is the necessary and sufficient cognitive condition for hope. It is not sufficient for hope that the subject think of Q (and ~ Q) as being possible (as against impossible), for belief or disbelief will negate hope even though the subject thinks that Q is possible (v impossible), and when unbelief is in place, though subjective possibility is formally necessary for hope, it is cognitively redundant because it is satisfied by the unbelief condition whenever the subject has the question of the existence of Q in mind.

I argued further that to distinguish the "possible/probable" from the "actual" as well as from the "impossible/certain" requires analysis in terms of the concept of affirmation, and that "A believes that Q" is to be understood as "A affirms Q" in the analysis of "A hopes that Q". I scrutinised Day's claim that "A hopes that Q" is to be analysed in terms of subjective probability rather than subjective possibility, which rests on his thesis that "A believes that Q" → "A estimates the subjective probability of Q as >1/2<1" and rejected this on the grounds that such a theory involves reducing all the modalities necessary to understand the concept of hope and belief to that of evidential justification. This cannot be done coherently and commits the fallacy of reducing the existence of an object to the conditions under which it can be known to exist.

I then applied the concept of hope developed to Kant's moral argument for the existence of God, and argued that it is sound if its conclusion is that commitment to morality requires agents to hope but not believe or disbelieve

---

Kant continues that to make the supreme good (happiness standing in exact relation with morality) complete "he who behaves in such a manner as not to be unworthy of happiness must be able to *hope* [my emphasis] that he will participate in happiness" [18, A813 B841].

If Kant sticks to hope and leaves belief (as normally understood) out of it (so that to postulate God's existence is not to assume that God exists but to assume only the possibility that God exists, i.e., to hope that God exists), then we have in this earlier version of his argument for God what seems to me the basis of a sound argument for necessity of hope for God.

that God exists. This has what some might consider a rather startling consequence, which derives from the fact that if God exists, since the moral law must be God's law, God must want us to eschew belief in God, yet hope that God exists.

This being so, I should make it clear that I am not overly concerned about just how faithful my reconstruction or interpretation of Kant's argument is to his intentions. What I claim is that, whatever Kant really intended, the interpretation or reconstruction I have presented renders Kant's position impervious to his moral attack on a supposed proof of theism, simply because his practical theism does not then involve propositional belief that God exists. And this, furthermore, fits his adamant assertion that practical theism cannot provide a premise for knowledge claims [20, 5: 137]. Above all, it has the advantage of rendering his position sound!

REFERENCES

[1] Aquinas, T. (2006).*Summa Theologiae: Fear and Anger (Volume 21: 1a2ae. 40-48).* Fathers of the English Dominican Province (Trans.) and J. P. Reid (Ed.). Cambridge, UK: Cambridge University Press.

[2] Aristotle, (1931). De Memoria et Reminiscentia. In J. I. Beare (Trans.) and W. D. Ross (Ed.), *The Works of Aristotle translated into English,* III. Oxford, UK: Clarendon Press.

[3] Beck, L.W. (1960). *A Commentary on Kant's Critique of Practical Reason*. Chicago, IL: University of Chicago Press.

[4] Beyleveld, D. (2009). *Morality and the God of Reason*. Inaugural Lecture at Utrecht, Netherlands: University of Utrecht.

[5] Beyleveld, D., and Brownsword, R., (2001). *Human Dignity in Bioethics and Biolaw*. Oxford, UK: Oxford University Press.

[6] Beyleveld, D., and Brownsword, R., (2010). Complex Technology, Complex Calculations: Uses and Abuses of Precautionary Reasoning in Law. In P. Sollie and M. Duewell et al (Eds.), *Evaluating New Technologies. Methodological Problems for the Ethical Assessment of Technological Developments* (pp.175-190). Library of Ethics and Applied Philosophy. Berlin, Germany: Springer.

[7] Bhaskar, R. (1975). *A Realist Theory of Science.* Sussex, UK: Harvester Press.

[8] Day, J. P. (1991). *Hope: A Philosophical Inquiry*. Helsinki, Finland: Acta Philosophica Fennica.

[9] Downie, R. S. (1963). Hope. *Philosophy and Phenomenological Research*, 24, pp.248-253.
[10] Fromm, E. (1974). *The Anatomy of Human Destructiveness.* London, UK: Routledge.
[11] Godfrey, J. J. (1987). *A Philosophy of Human Hope*. Dordrecht, Netherlands: Nijhoff.
[12] Harrison, J. (1963). Symposium: Christian Virtues. *Proceedings of the Aristotelian Society Supplementary Volumes*, 37, pp.69-82.
[13] Heil, J. (1982). Seeing is Believing. *American Philosophical Quarterly*, 19, pp.229-239.
[14] Hume. (1739). *A Treatise of Human Nature*. L. A. Selby-Bigge (Ed.) (1888), Oxford, UK: Clarendon Press.
[15] Marcel, G. (1951). *Homo Viator*. E. Crawford (Trans.). London, UK: Gollancz.
[16] May, R. (1970). *Love and Will*, London, UK: Souvenir Press.
[17] Mill, J. S. (1885). *Nature, the Utility of Religion, and Theism* (3rd edn). London, UK: Longmans Green.
[18] Kant, I. (1781, 1st edn A; 1787 2nd edn B). *Critique of Pure Reason.* N. Kemp Smith (Trans.) (1968). London, UK: Macmillan.
[19] Kant, I. (1785). *Groundwork of the Metaphysics of Morals*. M.Gregor (Ed.) (1998). Cambridge, UK: Cambridge University Press.
[20] Kant, I. (1788). *Critique of Practical Reason.* M. Gregor (Ed.) (1997). Cambridge, UK: Cambridge University Press.
[21] Kant, I. (1797). *The Metaphysics of Morals.* Mary Gregor (Trans.) (1991). Cambridge, UK: Cambridge University Press.
[22] Kant, I. (1790). *Critique of the Power of Judgment.* P.Guyer (Ed.) (2000). Cambridge, UK: Cambridge University Press.
[23] Leibniz, W. G. (1934). *Philosophical Writings.* M. Morris (Trans.). With an introduction by C.R. Morris. London, UK: Dent Dutton.
[24] Phillips-Griffiths, A. (1990). Certain Hope. *Religious Studies*, 26, pp.453-461.
[25] Price, H. H. (1969). *Belief.* London, UK: Allen and Unwin.
[26] Spinoza. (1955). *Ethics.* In R. H. M. Elwes (Trans. and Ed), *Works of Spinoza (Vol II).* New York, NY: Dover.
[27] Stotland, E. (1969). *The Psychology of Hope: An Integration of Experimental, Clinical and Social Approaches.* San Francisco, CA: Josey-Bass.

[28] Sutherland, S. R. (1989). Hope. In G. Vesey (Ed.), *The Philosophy in Christianity (*pp. 193-206). Cambridge, UK: Cambridge University Press.

[29] Voltaire, F-M. A, de. (1759). *Candide and Other Tales.* Tobias Smollett (Trans.), revised J. C. Thornton (1971). London, UK: Dent Dutton.

[30] Wheatley, J. M. O. (1958). Wishing and Hoping. *Analysis*, 18, pp.121-131.

In: Philosophy of Mind ISBN: 978-1-62257-215-1
Editors: R. J. Jenkins, W. E. Sullivan

Chapter 2

HOW CONSCIOUSNESS BUILDS THE SUBJECT THROUGH RELATING

Giorgio Marchetti*
www.mind-consciousness-language.com research center
via Guido Rossa, 8 Pozzo d'Adda, (Mi) Italy

ABSTRACT

This paper aims to show that the main difference that consciousness makes to human behavior is to provide us with a sense of self. Consciousness does this by allowing us to relate ourselves to other entities, and therefore to understand what kinds of relations exist between us and them. Variations in the state of nervous energy elicited by the use of attention are the basic underlying mechanism of consciousness. They are used to put things in relation, mainly by acting as the basis for the construction of possible orders (such as space and time).

Keywords: Consciousness, sense of self, relations, attention, nervous energy, order

* Correspondence E-mail: info@mind-consciousness-language.com.

INTRODUCTION

The present-day philosophical debate and scientific research on human consciousness pose a fundamental question: Does consciousness epiphenomenally accompany our brain activity, with no consequences on our thoughts, actions, perceptions, etc. or does it play a role in our behaviour? And if so, what is its function?

One could certainly be led to accept the idea that consciousness has no significant function, after considering the bulk of empirical evidence showing that human beings process a large amount of information unconsciously (Merikle et al., 2001), that complex decisions are best made after a period of distraction assumed to elicit unconscious thought (Dijksterhuis and Nordgren, 2006; Zhong et al., 2008), and that a freely voluntary act is not initiated by the subject's conscious free, but by his brain's unconscious processes (Haggard, 1999; Haggard and Eimer, 1999; Haggard et al., 1999; Libet, 1985; Libet et al., 1983).

This idea is further supported by the fact that very sophisticated information processing, such as recognizing faces, analyzing speech and chess playing, can be performed not only by human beings, but also by modern computers, which few of us would qualify as conscious agents.

Some scholars, such as Rosenthal (2008) for example, who argues that "the consciousness of thoughts, desires, and volitions adds little if any benefit for rational thinking, intentional action, executive function, or complex reasoning" (*ibid.*, p. 839), have concluded that consciousness has no significant function.

However, if the available data is examined more carefully, things are not that simple.

My aim in this paper is to show: that consciousness does make a difference to human behaviour; what this difference is; how such a difference comes about in and through consciousness; the underlying mechanism of consciousness.

Most of the ideas presented here can be found in Marchetti (2010), which contains a more detailed discussion, as well as more extensive reference to supporting empirical evidence.

CONSCIOUSNESS PLAYS A ROLE IN OUR BEHAVIOUR

Does consciousness make any difference to human behaviour? Contrary to those who maintain that consciousness is a mere epiphenomenon with no consequences on a person's behaviour, there are many reasons to believe that consciousness *does* play a role. Four such reasons are listed here below.

Firstly, studies using different experimental paradigms show that conscious and unconscious processing can lead to qualitatively different consequences. In Pavlovian conditioning studies for example, Clark and Squire (1998) show that trace conditioning requires an awareness of the conditioned stimulus-unconditioned stimulus relationship for conditional response acquisition, whereas awareness does not appear to be necessary for simple delay conditioning[1].

In implicit learning studies, Fu et al. (2008) – consolidating the findings made by Destrebecqz and Cleeremans (2001) and confirming predictions by Cleeremans and Jiménez (2002) - show in their Experiment 3 that unconscious (versus conscious) knowledge arises early in training. Unconscious knowledge is characterized by weak, poor-quality representations that are already capable of influencing a person's performance and are too weak for the person to be able to exert control over them[2].

Secondly, the interpretation of empirical data supporting the primacy of unconscious processing is not as straightforward as it may initially seem.

Rey et al.'s (2009) experiment clarifies the claim made by Dijksterhuis et al. (2006) about the supremacy of unconscious over conscious thought at solving complex decisions. By using an experimental design similar to the one used by to Dijksterhuis et al. (2006) but with an additional control condition (the "immediate condition") in which subjects made their choice immediately without any period of thought (conscious or unconscious), Rey et al. showed that decisions made by subjects in the immediate condition were just as good as those in the unconscious one, hence challenging Dijksterhuis et al.'s (2006) interpretation.

The same finding was replicated by Waroquier et al. (2010). Additionally, they found that while too much conscious deliberation can actually deteriorate

[1] Bekinschtein et al. (2011) observe that at least three cases seem to challenge this view: trace conditioning can be 1) learnt by almost every animal, even sea slugs; 2) elicited using subliminal stimuli; 3) learnt by clinically-defined unconscious patients. However, as their analysis reveals, the three cases do not contradict the possibility that consciousness is a prerequisite for trace conditioning.

[2] For further examples in other field studies, see Merikle and Daneman (1998).

high-quality first impressions, conscious thought enhances the quality of decisions in the absence of such prior first impressions[3].

In the research field of unconscious priming, Kunde et al. (2003) show that - contrary to the hypothesis that unconscious priming originates from purely unconscious semantic processing of the prime - unconscious primes activate responses to the degree that they match pre-specified action-trigger conditions: that is, the impact of subliminal stimuli is crucially determined by the subject's pre-stimulus intentions.

Likewise, Libet's (2004) conclusions about his experiments seem unjustified. As I have tried to show (Marchetti, 2005), in his experiments subjects *were aware* of the task they had to accomplish *well before* the time in which the freely voluntary act was to occur. Therefore, in a causal chain of events, a conscious decision to perform a freely voluntary act indisputably precedes the act itself.

Some researchers are also very critical towards empirical evidence (supported by implicit learning studies: Destrebecqz and Cleeremans, 2001; Nisbett and Wilson, 1977; Reber, 1967) proving that learning can be dissociated from awareness. In reviewing literature on implicit learning, Shanks (2005) finds that "it has yet to be proved beyond reasonable doubt that there exists a form of learning that proceeds both unintentionally and unconsciously" *(ibid.*, p. 216), and Perruchet (2008) argues: "there are quite limited supports to claim that while they perform the implicit test participants (1) have no conscious knowledge about the study material, (2) have the subjective experience of guessing, or (3) have no control over the expression of their knowledge" *(ibid.*, p. 615).

Thirdly, studies such as those by Perruchet and Vinter (2002), show that, in a dynamic perspective, conscious mental life is sufficient to account for (at least part of) human behaviour without any need to resort to the concepts of unconscious representations and knowledge, and the notion of unconscious inferences. The "mentalistic" framework put forward by Perruchet and Vinter, and expressed by the concept of self-organizing consciousness (SOC), proves to be more parsimonious than the prevailing view of the mind being based on the postulate of an omnipotent cognitive unconscious. Their model adopts

[3] Moreover, it should be noted that Dijksterhuis and Nordgren themselves cannot fail to recognize the usefulness of consciousness in aspects of human information processing at least: while arguing that in many ways unconscious thought is superior to conscious thought, they admit that "this superiority of unconscious processes does not pertain to the earlier stage of information acquisition. At that stage, conscious processes are superior" (Dijksterhuis and Nordgren , 2006, p. 106).

basic associative principles that operate on successive conscious contents while respecting and taking advantage of the constraints inherent to the conscious system, such as limited capacity, seriality and relative slowness of processing and memory decay. As such, it appears to be capable of generating highly complex representations that are able to fulfil functions which are generally assigned to unconscious rule-governed thinking.

Fourthly, experiments directly comparing the brain activation evoked by conscious versus nonconscious stimuli, have revealed that conscious processing involves neural processes other than those involved in nonconscious processing, thus supporting – albeit indirectly - the idea of a distinct functional role for consciousness[4]. For example, time-resolved experiments using ERPs aimed at following the processing of a visual stimulus in time as it crosses the threshold for conscious perception, show that (i) during the first 250 ms subliminal stimuli produce a transient, small and brief activation that progresses from the occipital pole toward both parietal and ventral temporal sites (ii) at around 270 ms conscious stimuli elicit a sudden onset of high-amplitude activity that is broadly distributed in the inferior and anterior prefrontal cortex as well as in the posterior parietal and ventral occipito-temporal cortices (Dehaene, 2009; Del Cul et al., 2007).

Accordign to Lamme (2003, 2006, 2010), a similar division of stimulus processing in different stages is associated with different levels of conscious access but is based on different neural mechanisms. Rees et al. (2002) offer a possible explanation that reconciles Lamme's and Dehaene's accounts.

EEG studies also provide evidence that different brain processes underlie conscious and nonconscious processing. Fingelkurts et al.'s (2011) study on vegetative (VS) and minimally conscious (MCS) patients shows that the size and duration of local EEG fields are smallest in VS patients, intermediate in MCS patients and highest in healthy fully conscious subjects. At the same time, these fields are quite stable in healthy subjects, less stable in MCS patients and very unstable in VS patients. The number and strength of coupling of local EEG fields (thought to be responsible for integrated subjective experiences) are highest in healthy subjects, intermediate in MCS patients and smallest or even absent in VS patients.

In short, although some contend that consciousness is a mere epiphenomenon that has no consequences on a person's behaviour, evidence shows that conscious processing does lead to behavioural consequences that

[4] However, it must be noted that there is no general consensus among researchers not only regarding the neural basis or correlate of consciousness, but also over the investigation stance and methods for determining the neural basis of consciousness (Lamme, 2010).

are qualitatively different from those produced by unconscious processing, and that different neuronal processes underpin conscious and unconscious processing.

Moreover, there are also strong arguments refuting the validity of evidence that certain phenomena, such as implicit learning and unconscious priming, can occur in the absence of awareness. This supports the view that part of human behaviour can be explained through conscious processing only.

FROM AN INFORMATION-PROCESSING TO A PERSON APPROACH TO CONSCIOUSNESS

I think that much of the misunderstanding concerning the role played by consciousness in a person's behavior originates from the level of analysis that is usually adopted by researchers when investigating it: namely, the information-processing conception of mind.

In its best known version, the information-processing approach considers the mind as a computer that processes information coming from external or internal sources in order to allow the person to provide appropriate behavioral responses. The information processed by the mind flows from one module to the next, until it reaches the last module of the chain: consciousness. Following the computer metaphor of the mind, this last module has been termed an "operating system" (Johnson-Laird, 1988), a "central processor" (Umiltà, 1988), or a "supervisory system" (Shallice, 1988).

The information-processing approach has certainly yielded various and positive results in psychological research on the mind: it can tell how long it takes for information to become conscious (Cleeremans and Sarrazin, 2007; Libet, 2004), the different levels of processing information involved by conscious vs. unconscious processes (Dehaene, 2009; Kouider and Dehaene, 2007), the different consequences that consciously vs. unconsciously processing information has on memory, learning, etc.

However, it is not the most appropriate approach when studying consciousness because it can neither provide the adequate level of analysis of the phenomenal aspect of consciousness, nor account for the emergence of the sense of self[5].

[5] By the expression "the sense of self" I mean the sense of a minimal self - with its sense of agency and ownership - and the sense of a narrative self (Gallagher, 2000; Hohwy, 2007),

The information-processing approach cannot provide the adequate level of analysis of the phenomenal aspect of consciousness simply because it does not address it. The information-processing approach analyzes the processes involved in transforming and elaborating information, the time needed to process information, how information is transformed, transmitted and disseminated, and so on, but not why these processes give rise to phenomenal experience. Let's consider for example Baars' (1988) Global Workspace model. Baars' model is certainly highly valuable in explaining a number of cognitive processes such as the subject's access to information, the influence of unconscious processors, voluntary control, reportability, etc. However, as Chalmers (1996) observes, the best that Baars' theory can do is to state that the information processed within the Global Workspace is experienced because it is globally accessible. But the question of why global accessibility should give rise to conscious experience remains unanswered. Not having directly and positively addressed the problem of the phenomenal aspect of consciousness, but rather having addressed derivative characteristics of conscious states (such as being "largely widespread and broadcast"), Baars' model can explain the latter, but not the former.

The information-processing approach cannot account for the emergence of the sense of self because its main concern is to analyze the *piece of information* processed by the person, or *how* this piece of information is transformed, rather than to analyze what it means for a *person* to consciously experience the piece of information that he/she is processing.

Adopting Negrotti's terminology (1997, 1999), we can say that the observation level of the information-processing approach is that of the *information* processed, or that of the *process* involved in processing information, not that of the *person* processing the information. As such, the information-processing approach cannot account for how a person (and his/her self) emerges, develops, changes and transforms by and through consciously processing information, but only for how some parts of a person's organism - sense organs, attention, memory, central processor, and so on - process information.

As various researchers (Cisek, 1999; Edelman, 1989; Freeman, 1999; Harnad, 1990; Searle, 1980, 1984, 1992) have highlighted, most of the problems raised by the information-processing approach are due to the fact that this approach considers information as made up of ready-made symbols

as well as of self-consciousness and the other possible manifestations of the sense of owning or being the subject of one's own conscious experiences.

representing the external world, whose meanings derive not so much from the history of the person, the importance they have for the person, his/her relations with other entities[6], but from the researcher's research goals.

Therefore, if we want to study consciousness, we have to change or perspective and no longer consider information, as well as the person processing it, as ready-made entities. On the contrary, we need to investigate how a person develops, changes and transforms by processing information, why and how something becomes "information" for a person, how something acquires a meaning for a person.

As we will see, and as I have tried to show elsewhere (Marchetti, 2010), by investigating this, we will find that both the person and information are the outcome of a continuous activity of differentiation carried out by an organism by applying its nervous energy to itself. In relating the organism to other entities, this activity allows the person and the other entities to co-emerge and come into existence.

The main tool that an organism possesses to carry out this activity of differentiation from other entities and emerge as a person, is consciousness. By making the organism directly experience how other entities relate to it (and how other entities relate to each other), consciousness is the privileged means by means of which an organism determines what relations exist between it and other entities, acquires a knowledge of itself and other entities, assigns a meaning to itself and other entities, and defines its own boundaries and shape, thus emerging as a differentiated entity: a person.

Summarizing, then, in order to fully account for the role that consciousness plays in a person's behavior, emergence and development, it is necessary to move the scientific analysis from an investigation of the difference that it makes for *information* to be consciously (vs. unconsciously) processed, to an investigation of the difference that it makes for a *person* to consciously (vs. unconsciously) process information.

THE SENSE OF SELF

An appeal to change the direction of scientific investigation on consciousness from an information-processing to a person approach has been

[6] By "other entities" I mean not only the objects, beings, organisms and events of the environment, but also the products of the very organism's activities, such as its movements, thoughts, plans, etc.

made by Cleeremans (2008, 2011)[7]. Cleeremans answers the question about what the notion of conscious subjective experience or *quale* means, by rejecting proposals such as Tononi's (2007), which, in analyzing conscious experience as a rather abstract dimension or aspect of information, would seem to overlook fundamental facts about experience. "Experience – *what it feels like*" observes Cleeremans "is anything but abstract. On the contrary, what we mean when we say that seeing a patch of red elicits an 'experience' is that the seeing *does something to us*" (Cleeremans, 2008, p. 20).

Cleeremans' proposal about what conscious subjective experience does to us, includes three main elements: memory, emotion and the sense of being the subject of one's experiences. In my view, the first two elements - memory and emotion -, while certainly identifying some important aspects, are not strictly necessary to qualify conscious subjective experience. Firstly, these two characteristics can also be elicited by unconscious processing. For example, emotional responses such as fear can occur without any awareness of their triggering stimuli (Tsuchiya and Adolphs, 2007). Secondly, these two characteristics seem to describe byproducts or second-order, albeit important, effects of conscious subjective experience, instead of grasping the fundamental differences that conscious (vs. unconscious) experience makes to a person and that characterize all conscious experiences. Not all our conscious experiences are accompanied by emotion or memories, which, on the contrary and most of the times, are consequent upon and triggered by a given conscious experience. Thirdly, there is no reason why some other equally important aspects or processes that usually characterize subjective experience should be ruled out. Examples of these are perceptions and thought, as well as all those processes listed by Baars (1988), or Seth (2009).

The third element identified by Cleeremans - the sense of being the subject of one's experiences - seems to be more plausible however: "it does not make any sense to speak of experience without an *experiencer* who experiences the experiences" (*ibid.*, p. 21); a "thermostat fails to be conscious because, despite the fact that it can find itself in different internal states, it lacks the ability to remove itself from the causal chain in which it is embedded. In other words, it lacks knowledge *that* it can find itself in different states; (…) there is no one home to be the *subject* of these experiences" (*ibid.*,

[7] Some other researchers have also proposed, each in their own way, changing approach and adopting a view that is more centered on the self and the person: see, for example, Damasio (1999, 2000), Evans' (1970) proposal of the "self-approach", Metzinger's (2000) "phenomenal self-model", Varela's (1996, 1999) and Varela and Shear's (1999) proposal to reconcile the first- and the third-person approach to consciousness.

p. 21). All of our conscious experiences are indeed accompanied – either at the time we have them or at a later stage - by the feeling that they belong to us, in the sense that we feel that it is we, and not someone else, who are experiencing them (Thompson, 2007).

This feeling, which I term the sense of self, is continuously brought about and reinforced (Fingelkurts and Fingelkurts, 2011) every time: we (decide/are able to) recall our past conscious experiences and experience them again; we decide to stop having a given conscious experience and start having another one; we realize that a conscious experience modifies and changes as we modify and change (for example, our conscious experience of duration changes with our age, psychological state, the kind of activity we are performing, etc.: see Flaherty, 1999; Marchetti, 2009); we observe that the conscious experience we are having is a precise and specific one, which differs from other ones.

The sense of self is dynamic, not singular but multiple, emerging chronologically in development "like onions, layers after layers, in a cumulative consolidation" (Rochat, 2003, p. 730).

Therefore, one of the main differences when a person consciously (vs. unconsciously) processes information, is definitely that it provides him/her with the sense of self. How does this sense of self affect and characterize a person's conscious experiences?

Firstly, a person's conscious experiences are characterized by the sense that they primarily originate from and refer to him/herself: what the person feels depends on him/her. This is quite patent when the person's actions and thoughts are involved. Yet it is also evident with other kinds of conscious experiences: there is no perception which does not originate from or is related to the person's specific sense-organs, body, position in space, way of looking at and conceiving the world, etc.

Secondly, it makes a person feel that the conscious experiences he/she has, have a direct effect on him/her, in the sense that every conscious experience directly affects, changes, modifies or transforms him/her: what a person feels has an effect on him/her. The most obvious examples of these effects are extreme and acute sensations such as pain, effort, exhaustion, pleasure, gratification, relief, thirst, hunger, etc., but there are also countless examples in the various sensory domains represented by sensations having intermediate or minor effects.

Thirdly, it provides a person with the sense of being a persistent, coherent entity, self or agent. Even if a person undergoes changes and various experiences, he/she continues to exist as a unified and coherent whole, on and

from which all his/her actions and experiences center and evolve: what the person feels makes him/her experience to be a consistent, persistent, unique agent.

Fourthly, it makes a person understand that he/she can directly control and guide the course of his/her own actions by means of his/her conscious activity. One a person has understood this – or, paraphrasing Cleeremans (2008), once he/she has "learnt to be conscious" -, he/she equips him/herself with the capacity to regulate him/herself. This means setting his/her own aims and objectives, taking decisions, evaluating events and situations, learning new strategies, adapting to changes, etc. Simply put, being self-conscious. From that moment on, a person's actions are primarily and directly governed not so much by innate instincts as by what happens in the person's consciousness, even though the latter can be occasioned by the former.

According to Vogeley et al. (2004) and Schilbach et al. (2008), the neurophysiological basis of self-consciousness or the sense of "being a self" is provided by the default-mode network (DMN). Fingelkurts and Fingelkurts' work (2011) brings further support to this hypothesis. By showing that the integrity of DMN persists unchanged across a variety of different cognitive tasks - and therefore is task-unrelated -, they can account for the fact that a subject that experiences phenomenal self-consciousness *always* feels directly present in the centre of a multimodal perceptual reality.

THE MAIN ACTIVITY THAT CONSCIOUSNESS ALLOWS A PERSON TO DO

But what are the conditions necessary for the sense of self to emerge and take shape? A very plausible answer comes from developmental psychology when it describes how the subject forms and develops. As Piaget (1974, pp. 281-282) has suggested, the subject (in my terminology, the person) only learns to know himself when acting on the object (in my terminology, other entities), and the latter can become known only as a result of progress of the actions carried out on it. In other words, by continuously acting, the subject *differentiates* him-herself from the object and emerges as a cognizing agent: this process allows the subject to define him-herself and its own boundaries while also defining the boundaries of the object. It should be noted here that "object" refers to both inanimate and animate entities. In fact, according to various researchers, in order to emerge, a person's higher-order cognitive

processes, such as those involving meta-representational self-consciousness (Newen and Vogeley, 2003), require a person to interact with other persons, so that the person can compare and distinguish his/her own cognitive states with those of others (Decety and Chaminade, 2003; Newen and Vogeley, 2003; Schilbach et al., 2008).

The process of differentiation which allows the subject to emerge is based on a very general activity: the possibility for the subject *to relate* to other entities. This activity includes not only very basic and simple ways of relating - such as the same/different distinction, symmetry, repetition, etc. - but also more complex and abstract ones. In the differentiation process, the person comes to learn and understand: how, when, where and why he/she relates with other entities; the value and functions that other entities have for him/her; how the person can affect or change other entities; how other entities affect him/her and make him/her change or not change. Subsequently, and on the basis of this first-level knowledge of the relations between him/her and other entities, the person can build a second-level knowledge of the relations existing between other entities.

In my view, this is precisely the main activity that consciousness allows the person to perform: that is, *it allows the person to relate him-herself to other entities* and therefore to understand what the relations between him/her and other entities. Consciousness, making the person experience directly what he/she is doing, the results of his/her activity, how he/she can affect other entities (and vice versa), how other entities limit him/her, etc., is the privileged way a person has of recognizing the relation between him-herself and objects and therefore of defining him-herself and other entities.

Some could argue that because animals can also perform elementary tasks of relational learning, and because it is questionable whether animals have any form of consciousness, consciousness is not necessary in order to (learn how to) place entities in relation to each other.

Apart from the plausibility of the claim that animals do not possess any form of consciousness (see Northoff and Panksepp, 2008; Panksepp, 2005), the argument that consciousness is not necessary to learn and place entities in relation to each other can be refuted on the grounds that there is abundant evidence showing the opposite, at least as far as more complex forms of relations are involved.

Firstly, Sackur and Dehaene (2009, experiment 4) have shown that while simple tasks - such as performing an arithmetical operation on a target number or comparing the target number to a reference number – can be performed even when the target number is masked and cannot be consciously perceived,

a chained (or composite) task made of two simple tasks on one target number cannot be performed in the absence of consciousness.

Secondly, classical conditioning studies show that "awareness is a prerequisite for successful trace conditioning" (Clark and Squire, 1998, p. 79). According to Clark and Squire, the more complex condition involved in trace conditioning vs. a simpler form of conditioning such as delay conditioning, would require consciousness to represent and remember the temporal conditioned stimulus-unconditioned stimulus relationship (for similar considerations in a Pavlovian conditioning study but involving a *detection* task vs. an *identification* task, Núñez and de Vincente, 2004).

Thirdly, in fear conditioning studies, a similar view on the role of consciousness has been expressed by Knight et al. (2006, p. 160), who found that awareness is necessary for conditional responding during trace, but not delay, fear conditioning (see also Carter et al., 2003).

Fourthly, the psychology of perception shows that the order of perceived events is highly dependent on whether or not their duration falls inside the "phenomenal present" (*Präsenzzeit*) (Stern, 1897), that is, the interval of physical time that, despite being composed of non-contemporaneous parts, is perceived as a unitary and unique act of consciousness (for a discussion, see Fingelkurts et al., 2010). As Vicario (2005) extensively shows, when all the single phases of a sequence fall into the phenomenal present, the sequence can undergo some kind of restructuring (according to certain Gestalt principles of organization) irrespective of the physical temporal contiguity of the stimuli. On the contrary, if the single phases of a sequence occupy a whole phenomenal present, the sequence of stimuli cannot undergo any kind of restructuring and the sequence of the perceived stimuli will be the same as the sequence of physical stimuli. The phenomenon of "temporal displacement" investigated by Vicario (1963), as well as other phenomena such as "continuous displacement", "tunnel effect", "Renard effect", "window effect" and "phi phenomenon", show that how relations between objects and events appear to us is strongly determined by the very features of the conscious working (such as the duration of the phenomenal present), to the extent that the order of perceived events does not correspond to that of the physical events.

However, the clearest evidence that a person must be conscious in order to create and place entities in relation to each other, is definitely represented by the extensive creation and use of natural and formal languages by human beings (Chafe, 1994). This is a prerogative which does not belong to other

species that posses simpler forms of consciousness or do not possess consciousness at all.

Artificial and natural languages provide a wide and specialized variety of ways of connecting and correlating real and abstract objects and events, as exemplified by logical connectives, mathematical operators and "grammatical" words and morphemes (Benedetti, 2009, 2010, Ceccato and Zonta, 1980). The fact that language represents a unique and specialized tool in connecting objects and events is further exemplified by the evidence reported by Conway and Christiansen (2001) of the strong connection between language and the ability to encode and represent the order of discrete elements occurring in a sequence (sequential learning).

The hypothesis that the main feature of consciousness is to allow a person to (learn how to) place entities in relation to each other, can also be criticized by saying that it is too restrictive and does not account for everything that consciousness allows a person to do.

Undeniably, such a criticism would seem more than reasonable if one considers, for instance, the eighteen functions listed by Baars (1988). The variety of these functions however can be reduced to the one I propose inasmuch as they let the person relate him-herself to other entities.

Summarizing then, what makes the sense of self possible is the fact that a person can, by means of his/her consciousness, place entities in relation to each other, and that, by means of placing entities in relation, he/she can differentiate him-herself from other entities. But how does this activity of placing things in relations with each other through consciousness occur? What mechanism allows consciousness to do this? More in general, what is the underlying mechanism of consciousness?

THE UNDERLYING MECHANISM OF CONSCIOUSNESS

My main hypothesis is that consciousness is the result of a person's attentional activity - that is, the continuous use and application of his/her attention - and that through his/her attentional activity, a person understands what relations exist between him/her and other entities, what his/her own boundaries and limits are, and therefore the limits and boundaries of other entities.

The hypothesis is based on six fundamental tenets, which I will describe here briefly. The reader can find a detailed presentation and discussion of supporting empirical evidence in Marchetti (2010).

Firstly, attention is the core element necessary, even if not sufficient, for consciousness: without attentional activity, there cannot be consciousness. The position about whether attention is necessary for consciousness ranges from those who maintain that attention and consciousness are distinct phenomena that need not occur together (Koch and Tsuchiya, 2006, Lamme, 2003) to those who maintain that the two are inextricably linked (De Brigard and Prinz, 2010, Mack and Rock, 1998, Posner, 1994). As I have tried to show (Marchetti, 2010), the view that in general there can be consciousness without attention originates primarily from the failure to notice the varieties of forms and levels of attention (Chun et al., 2011; Demeyere and Humphreys, 2007; La Berge, 1995; Lavie, 1995; Nakayama and Mackeben, 1989; Pashler, 1998) and consciousness (Bartolomeo, 2008; Edelman, 1989; Iwasaki, 1993; Vandekerckhove and Panksepp, 2009). Not all forms of attention produce the same kind of consciousness, and vice versa not all forms of consciousness are produced by the same kind of attention. Overlooking this fact may lead to the wrong view that there can be consciousness without any form of attention (for a similar criticism, see Kouider et al., 2010, Srinivasan, 2008).

Secondly, attentional activity can be performed because of the nervous energy that is supplied by the organ of attention. The concept of "nervous energy" implies the ideas that nervous energy is a pool that allows us to perform a certain kind of work, is limited, runs out, is replenished, and can be used in a flexible way. The concept of nervous energy – for which alternative terms, such as "psychic energy", "limited capacity processor", "resource", and "effort" have also been used – has been analyzed in various ways in relation to attention by many researchers and authors (Csikszentmihalyi, 1992; Kahneman, 1973; Mach, 1890; Wickens, 1984). More in general, the notion of energy is currently used and investigated in relation to brain activity (Laughlin, 2001, Laughlin and Sejnoiwski, 2003, Shulman et al., 2009a, 2009b). As to the "organ of attention", many scientists have started investigating its physical substrate (Crick 1994, Crick and Koch, 2003; Mesulam, 1990; La Berge, 1995; Posner 1990, 1995; Posner and Petersen 1990; Knudsen, 2007).

Thirdly, attentional activity consists in the continuous application of attention to the other organs (sense organs, the proprioceptive system, the interoceptive, system, the musculoskeletal system, and working memory) or to attention itself. This "continuous" working of attention can best be conceived as cyclical, that is, a repetition of successive acts of focalization each of which has a specific minimal and maximal duration. The hypothesis of the cyclical dynamics of attention (Buschman and Miller, 2010; Large and Jones, 1999;

VanRullen et al., 2007; Ward, 2003) can also be inferred from the observation that no one can possibly attend continuously to an object that does not change (James, 1890), or from the close correlation between the perception of apparent simultaneity and the alpha phase at which stimuli are presented (Varela et al., 1981).

Fourthly, attentional activity allows a person to perform actions that can directly vary his/her own state of nervous energy. This variation constitutes the phenomenal aspect of consciousness, or qualia. I have partly derived this idea from Valéry's (1973) observation that sensation is a variation of the state of energy of a closed system. Other suggestions indicating that consciousness results from a variation of the organism's internal state can also be found in Damasio's (1999) work. The idea that consciousness arises as a consequence of the modification of the energetic state of the organ of attention induced by the use of attention itself, is partly captured by Haikonen (2003), who argues that some basic conscious states (pain pleasure, good and bad) derive from the modulation of attention.

Fifthly, by acting attentionally, a variation in the state of nervous energy is induced. This produces a conscious experience in the form of either a certain level of constraint or of non-constraints (to act in general). The constraints that a person experiences when acting attentionally are determined by the level of attention applied, the specific structure of his/her body and the relations resulting from the interaction between his/her body and other entities. This aspect was conveyed very well by Piaget (1937), who – in describing how the idea, or concept, of an object is built up during the first stages of intellectual development - observed that the subject recognizes his own reaction before he recognizes the object as such.

Sixthly, these constraints – which are the basic elements of conscious perception - consist precisely of the interruption, hindrance, slowing down, facilitation, stimulation, acceleration, and so on, of attentional activity. Whenever a person finds an obstacle or cannot extend his/her limbs beyond a certain length or cannot make a movement, his/her attentional activity, and all his/her being along with it, is slowed down or even temporarily stopped, so much so that the person must either apply his/her nervous energy in a new way or redirect it to something else, if he/she wants to unblock the situation.

A very interesting, albeit partial, exemplification of the mechanism by means of which attentional activity induces a variation in the state of nervous energy, is represented by Cabanac and Russek's (2000) model of regulated biological systems (for a discussion, see Marchetti, 2010). By offering a model of representation for both human consciousness and the other main

physiological functions (such as pulmonary ventilation, blood circulation, etc.), Cabanac and Russek also gives biological plausibility, from an evolutionary point of view, to my model of consciousness as being the evolution of more primitive systems.

Additionally, Cabanac (1996, 2003) puts forward the idea of the affective dimension of human consciousness as being the "common currency" for the trade-offs that take place in the mind to achieve a ranking of priorities. This is consistent with my idea that consciousness is the privileged way of controlling the other organs and systems by means of a unique and common form of energy: nervous energy[8].

The constraints a person experiences every time he/she acts attentionally, represent the basic elements that allow the person to come to know and define him-herself, other entities, and the relations between him-herself and other entities. For instance, the activity a person performs when trying to reach something unsuccessfully has a direct effect on him-herself, in that it modulates his/her own pool of nervous energy by either blocking the nervous energy flow, re-directing its course, or further stimulating it in the same direction. This effect, which constitutes a person's conscious experience, immediately gives this person the dimension of his/her effort, helps define the boundaries of both his/her body and the entity he/she tries to reach, and the relations between him-herself and the entity. The information so acquired constantly updates what I have called the "schema of self" in my model of consciousness (Marchetti, 2001, 2010). The schema of self is the system that incorporates and coordinates all the innate and learned values and schemata needed to keep the organism alive, and that provides all the rules which make our organism perceive, move, act in general and interact with other organisms.

On the basis of the initial experiences of his/her own boundaries and those of other entities, a person can subsequently refine the definition of the relations existing between him/her and other entities. As demonstrated by Piaget's (1936, 1937) influential work on developmental psychology, a person's relations are defined through a gradual and continuous process of differentiation and systematization. This is made of different stages that progressively generate and consolidate, among other things, the notions of independent entities (such as "object", "agent", "space", and "time").

As observed by Piaget (1936), the process leading to the definition of a person's relations with other entities requires some inborn behaviors, such as

[8] This aspect is also reminiscent of Ukhtomsky's (1966) principle of "dominanta": for a discussion of Ukhtomsky's principle, see Fingelkurts et al. (2010).

sucking and grasping reflexes, in order to be started. Additionally, it also requires some innately specified processes and principles that help infants direct and coordinate their attention to relevant aspect of the input.

In this regard, developmental psychologists have put forward various, and sometimes, contrasting, hypotheses (Karmiloff-Smith, 1992). For example, Piaget (1936, 1937) granted the newborn child three domain-general functional processes (assimilation, accommodation and equilibration). Spelke (1990) identified a set of specific principles that serves to guide their perceptual analysis of the physical word: boundedness, cohesion, rigidity and no action at a distance. Mandler (2008, 2010) goes so far as to hypothesize the existence of an innate attentional mechanism (which she calls Perceptual Meaning Analysis, PMA) that records selected aspects of incoming spatial information into an accessible conceptual format. During the first year of life, preverbal infants are able to reduce and redescribe perceptual information into a spatial image-schematic form thanks to the work done by PMA.

Social neuroscientists have also highlighted the importance of specialized neuronal systems that compute socially relevant stimuli (faces, gaze direction, motion of bodies and limbs) in defining a person's relations with other entities. Particularly interesting is Graziano and Kastner's (2011) proposal about the central role they assign to social attention (Birmingham and Kingstone, 2009; Frischen et al., 2007; Nummenmaa and Calder, 2008) in the construction of perceptual models of other minds and one's own mind. According to them, we are endowed with a social perceptual machinery which allows for the predictive modeling of the behavior of individuals. The existence of this machinery is supported by a series of empirical studies (Blakemore et al., 2003; Calder et al., 2002; Gallagher et al. 2000; Grossman et al., 2000; Pelphrey et. al., 2004, 2005, Saxe and Kanwisher, 2003; Saxe and Wexler, 2005; Vogeley et al. 2001). Its main component is the ability to track someone's attention, because it provides critical information for predicting an individual's behavior. Graziano and Kastner identify the superior temporal sulcus (STS) and the right temporoparietal junction (TPJ) as the main areas responsible for building perceptual models of minds. These two areas operate in a cooperative fashion with other areas and networks devoted to social perception: for example, the models generated by STS and TPJ would seem to be used by the mirror-neuron system (Rizzolatti and Sinigalia, 2010) to drive simulations.

It is important to note that while all these hypothetical processes and machineries can explain, for example, a person's capacity to assign the cause of experiencing a certain constraint to an external agent rather than to him-

herself, they cannot explain the experience of constraint itself, that is, how conscious phenomena and experiences generate.

My model of consciousness is not the only one having attention as one of its fundamental components. Other models also include (Baars, 1988; Chella, 2007; Graziano and Kastner, 2011; Haikonen, 2003; Madl et al., 2011; Taylor, 2002, 2007a, 2007b). However, none of these base their explanation of the phenomenal aspect of consciousness on attention (the only partial exception is Haikonen, 2003). In these models, attention is usually a privileged way to access consciousness, or a process that provides the needed information for consciousness, but it does not have any role in generating qualia. For example, for Graziano and Kastner (2011), consciousness is an (imprecise but useful) informational and perceptual representation of the process of attention. Therefore, in order to account for the phenomenal aspect of consciousness, these models need to resort to some additional process (such as perception or representation), thus pushing the explanation back, in an endless regression (how does perception or representation produce consciousness?). On the contrary, my model explains consciousness by means of a very basic process such as attention, without resorting to additional processes.

CONSCIOUSNESS CREATES ORDERS

In my view, the principal way in which the variations in the state of nervous energy (induced by attentional activity) are used to place entities in relation to each other, is by serving as a basis for the construction of possible orders (by "order" I mean all kinds of series, successions, arrangements, sequences, schemas and organizations in general).

Space and time are some typical examples. Once an order is created, it is possible to build various kinds of relations on this: spatial, temporal, causal, logical, physical, etc.

According to my analysis, an order can be created from variations in the state of nervous energy by exploiting some of the characteristics of the organ of attention and working memory, respectively, by applying attention to something in a continuous, incremental and cumulative way (to a certain extent at least), and by keeping track of the results of the work of attention.

In spatial order, the possibility of using working memory to keep track of the various conscious sensations produced by (applying attention to) a moving sense-organ, is fundamental in my opinion. Imagine you are perceiving a

surface by moving your index finger on it. By keeping present in consciousness, in an incremental way, the single sensations that are produced by the index finger while it moves, you create a sequence or succession of sensations, which is the basis for the formation of two-dimensional constructs, such as "path", "line" and "distance". As a result, by exploiting the various features of attention[9], more complex spatial constructions can be developed (for an analysis of how temporal order can be obtained by using the work of attention, see Marchetti, 2009).

Other researchers base their explanation of the origin of spatial experience on the operations of attention. According to Mandler (2008, 2010), PMA produces a set of spatial primitives, such as PATH, LOCATION, MOTION TRANSFER, which is sufficient to account for the early conceptualizations that preverbal infants use to interpret objects and events. These early conceptualizations are important because they represent the core ontogenetic foundations on which later concepts are built and play a major role in determining the organization of the adult conceptual system.

Carstensen (2007, 2011) shows that selective attention plays a central role in the characterization of spatial relations, and that the representation of attentional aspects leads to the possibility of defining an ontological upper structure which covers both the spatial and temporal domain.

For Scheider and Kuhn (2011), the human being experiences the geometrical and topological structures of the environment by performing and comparing attentional steps. An attentional step is the actual movement of attention from focus *x* to *y*. Scheider and Kuhn put forward an operational model of constructive geometry grounded on primitives of the human attentional apparatus, which allows for the referencing and predication of geometrically relevant Gestalt phenomena in vista environment. In particular, it allows for detecting whether one focus of attention precedes another one (primitive perception of time), whether attention focuses on the same point-like feature, whether a given pair of foci is congruent to another pair, and whether a focus points between two others.

The fact that consciousness allows us to build orders does not rule out that non-conscious systems can also build some kind of order. Artificial neural networks can learn rich structured representations that capture abstract

[9] Primarily, that attention is cyclical in nature, which allows for the realization of mental cognitive operations of increasing complexity (Benedetti et al., 2010; Fingelkurts and Fingelkurts, 2001, 2005, 2006, Fingelkurts et al., 2010). On the advantages offered by discrete neural computations over a continuously evolving system, see also Buschman and Miller, 2010).

dimensions of a given task domain, as a result of merely being required to process exemplars of the domain (Cleeremans, 2011). However, there are at least two important differences between conscious and non-conscious systems. Firstly, while non-conscious systems produce orders that can be successfully used only in the context of performing the particular task for which they were trained, conscious systems produces orders that can be applied without restriction to any domain. Secondly, while non-conscious systems produce orders that cannot be easily modified, consciousness produces orders that can even be completely inverted (for example, we can conceive of time as being reversible).

The experiences of constraints supplies the firm ground on which the person can build orders that allow him/her to create various kinds of relations: spatial, temporal, etc. Consciousness makes it possible to create orders, and to create relations based on these orders. Without consciousness, neither orders nor the relations that can be built on them, would exist. A person's knowledge assumes the form that his/her consciousness allows it to assume. Everything a person knows, he/she comes to know in and through his/her consciousness. A person comes to know the world as it is because of his/her conscious experience. Conscious experience is the only level of reality that a person can directly access. All the other levels can be accessed only indirectly via the privileged medium of consciousness. Consequently, the world appears to the person as his/her consciousness lets him/her experience it and unavoidably bears the hallmark of his/her consciousness. Its qualities and characteristics are the qualities and characteristics of his/her consciousness. The world is ordered according to the principles established by and through a person's consciousness, and has the form that such principles let it assume.

CONCLUSION

This paper shows that consciousness makes a difference to human behaviour, determining and influencing what a person thinks, does, feels, etc. The main difference is that it provides the person with the sense of owning or being the subject of his/her conscious experience, what I have called "the sense of self". A person can have such a sense of self because, in and through consciousness, he/she can relate him-herself to other entities, and therefore understand what relations exist between him/her and other entities. The basic elements that allow a person to relate him-herself to other entities, and more in general to place entities in relation to each other, are the variations of his/her

own state of nervous energy induced by the application of his/her attention. These variations, which constitute the phenomenal aspect of consciousness, or qualia, produce a conscious experience in the form of either a constraint or non-constraint (to act in general). We have also seen that the principal way in which these variations are used to relate things, is by serving as a basis for the construction of possible orders (such as space and time).

Most of the ideas I have put forward here, and more in general my theory of consciousness should be treated as unproven until they are verified. I believe however that they are supported by a bulk of evidence, that they are congruent with many of the current theories and views of consciousness, and that their explicative power can account for many unresolved dilemmas.

ACKNOWLEDGMENTS

I am grateful to Dr. Giulio Benedetti and Drs. Andrew and Alexander Fingelkurts for their constant support and insightful comments, and to Wendy Piemonte for the English revision of the article.

REFERENCES

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Bartolomeo, P. (2008). Varieties of attention and of consciousness: evidence from neuropsychology *Psyche* 14, 1-33.

Bekinschtein, T. A., Peeters, M., Shalom, D., Sigman, M. (2011). Sea slugs, subliminal pictures and vegetative state patients: boundaries of consciousness in classical conditioning. *Frontiers in Psychology* 2 (337).

Benedetti, G. (2009). The Meaning of the Basic Elements of Language in Terms of Cognitive Operations: Operational Semantics. *Advanced Studies in Biology* 6, 255-305.

Benedetti, G. (2010). "A Semantics of the Fundamental Structural Elements of Language, Based on Cognitive Functions: Operational Semantics", in *Cognitive Psychology Perspectives*, eds. G. Salvati and V. Rabuano (New York: Nova Science Publishers) 1-61.

Benedetti, G., Marchetti, G., Fingelkurts, A. A., and Fingelkurts, A. A. (2010). Mind Operational Semantics and Brain Operational Architectonics: A Putative Correspondence. *The Open Neuroimaging Journal* 4, 53-69.

Birmingham, E. and Kingstone, A. (2009). Human Social Attention. A New Look at Past, Present, and Future Investigations. *Annals of the New York academy of Sciences* 1156, 118-140.

Blakemore, S. J., Boyer, P., Pachot-Clouard, M,, Meltzoff, A., Segebarth, C., and Decety, J. (2003). The detection of contingency and animacy from simple animations in the human brain. *Cerebral Cortex* 13, 837-844.

Buschman, T. J. and Miller, E. K. (2010). Shifting the spotlight of attention: evidence for discrete computations in cognition. *Frontiers in Human Neuroscience* 4, 1-8.

Cabanac, M. (1996). On the origin of consciousness, a postulate and its corollary. *Neuroscience and Biobehavioral Reviews* 20, 33-40.

Cabanac, M. (2003). *La cinquième influence ou la dialectique du plaisir*. Laval, Québec: Les Presses de Luniversité

Cabanac, M. and Russek, M. (2000). Regulated biological system. *Journal of Biological System* 8, 141-149.

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., and Young, A. W. (2002). Reading the mind from eye gaze. *Neuropsychologia* 40, 1129-1138.

Carstensen, K.-U. (2007). Spatio-temporal Ontologies and Attention. *Spatial Cognition and Computation* 7 (1), 13-32.

Carstensen, K.-U. (2011). Towards Cognitivist Ontologies. On the role of selective attention for upper ontologies. *Cognitive Processing*, DOI: 10.1007/s10339-011-0405-0.

Carter, R. M., Hofstötter, C., Tsuchiya, N. And Koch, C. (2003). Working memory and fear conditioning. *Proceedings of the National Academy of Sciences of the United States of America* 100 (3), 1399-1404.

Ceccato, S. and Zonta, B. (1980). *Linguaggio consapevolezza pensiero*. Milano: Feltrinelli.

Chafe, W.(1994). *Discourse, Consciousness, and Time. The Flow and Displacement of Conscious Experience in Speaking and Writing.* Chicago and London: The University of Chicago Press.

Chalmers, D. J. (1996). *The Conscious Mind. In Search of a Fundamental Theory*. New York: Oxford University Press.

Chella, A. (2007). "Towards Robot Conscious Perception", in *Artificial Consciousness*, eds. A. Chella and R. Manzotti (Exeter, UK: Imprint Academic) 124-140.

Chun, M. M., Golomb, J. D. and Turk-Browne, N. B. (2011). A Taxonomy of External and Internal Attention. *Annual Review of Psychology* 62, 73-101.

Cisek, P. (1999). Beyond the computer metaphor: behavior as interaction. *Journal of Consciousness Studies* 6, 125-142.

Clark,R.E. and Squire, L.R. (1998). Classical conditioning and brain systems: The role of awareness. *Science* 280, 77–81.

Cleeremans, A. (2008). Consciousness: the radical plasticity thesis. *Progress in Brain Research*, eds. R. Banerjee and B.K. Chakrabarti, 168, 19-33.

Cleeremans, A. (2011). The radical plasticity thesis: how the brain learns to conscious. *Frontiers in Psychology* 2, 1-12.

Cleeremans, A. and Jiménez, L. (2008). "Implicit learning and consciousness: a graded, dynamic perspective", in *Implicit learning and consciousness*, eds. R. M. French and A. Cleeremans (Hove, UK: Psychology Press), 1-40.

Cleeremans, A. And Sarrazin, J.-C. (2007). Time, action, and consciousness. *Human Movement Science* 26 (2), 180-202.

Conway, C. M. And Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences* 5, 539-546.

Crick, F. (1994). *The Astonishing Hypothesis.* New York: Simon and Schuster.

Crick, F. and Koch, C. (2003). A framework for consciousness. *Nature Neuroscience* 6: 119-126.

Csikszentmihalyi, M. (1992). Flow. *The Classic Work on how to achieve happiness*. London: Random House Group.

Decety, J. and Chaminade, T. (2003). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Consciousness and Cognition* 12, 577-596.

Damasio, A. (1999). *The Feeling of What Happens. Body and Emotions in the Making of Consciousness.* New York: Harcourt Brace and Co.

Damasio, A. (2000). "A Neurobiology for Consciousness", in *Neural Correlates of Consciousness. Empirical and Conceptual Questions*, ed. T. Metzinger (Cambridge. Massachusetts: The MIT Press) 111-120.

De Brigard, F., and Prinz, J. (2010). Attention and consciousness. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 51–59.

Dehaene, S. (2009). Conscious and Nonconscious Processes: Distinct Forms of Evidence Accumulation? *Séminaire Poincaré* XII, 89-114.

Del Cul, A., Baillet, S., and Dehaene, S. (2007). Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLoS Biol* 5(10), 2408-2423.

Demeyere, N. and Humphreys, G. (2007). Distributed and focused attention: Neuropsychological evidence for separate attentional mechanisms when counting and estimating. *Journal of Experimental Psychology: Human Perception and Performance* 33, 1076-1088.

Destrebecqz, A., and Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the Process Dissociation Procedure. *Psychodynamic Bulletin and Review* 8, 343-350.

Dijksterhuis, A. and Nordgren, L. F. (2006). A Theory of Unconscious Thought. *Perspectives on Psychological Science* 1, 95-109.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., and van Baaren, R. B. (2006). On making the right choice: the deliberation-without-attention effect. *Science* 311, 1005-1007.

Edelman, G. M. (1989). *The Remembered Present: a Biological Theory of Consciousness*. New York: Basic Books.

Evans, C.O. (1970). *The Subject of Consciousness*. London: George Allen and Unwin Ltd.

Fingelkurts, An. A., and Fingelkurts, Al. A. (2001). Operational architectonics of the human brain biopotential field: Towards solving the mind-brain problem. *Brain and Mind* 2, 261-296.

Fingelkurts, An. A., Fingelkurts, Al. A. (2005) "Mapping of Brain Operational Architectonics", in *Focus on Brain Mapping Research*, ed. F. J. Chen (Hauppauge, NY: Nova Science Publishers Inc.) 59-98.

Fingelkurts, An. A., and Fingelkurts, Al. A. (2006). Timing in cognition and EEG brain dynamics: discreteness versus continuity. *Cognitive Processing* 7, 135–162.

Fingelkurts, An. A., and Fingelkurts, Al. A. (2011). Persistent operational synchrony within brain default-mode network and self-processing operations in healthy subjects. *Brain and Cognition* 75, 79-90.

Fingelkurts, An. A., Fingelkurts, Al. A., Neves, C. F. H. (2010). Natural World Physical, Brain Operational, and Mind Phenomenal Space-Time. *Physics of Life Reviews* 7 (2), 195-249.

Fingelkurts An. A., Fingelkurts Al. A., Neves C. F. H., Bagnato S., Boccagni C., Galardi G. (2011) Operational architectonics of consciousness: EEG study in patients with severely injured brain. *Proceedings of the Conference "Toward a Science of Consciousness: Brain, Mind, Reality".* Stockholm, Sweden, May 3-7, 84.

Flaherty, M. (1999). *A watched pot: How we experience time*. New York: New York University Press.

Freeman, W. J. (1999). *How Brains make up their Minds*. London: Weidenfeld, Nicolson.

Frischen, C. K., Bayliss, A. P., Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin* 133, 694-724.

Fu, Q., Fu, X. and Dienes, Z. (2008). Implicit sequence learning and conscious awareness. *Consciousness and Cognition* 17, 185-202.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia* 38, 11-21.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences* 4, 14-21.

Graziano, M. S. A. and Kastner, S. (2011). Human consciousness and its relationship to social neuroscience: A novel hypothesis. *Cognitive Neuroscience* 2 (2), 98-133.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., and Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience* 12, 711-720.

Haggard, P. (1999). "Perceived timing of self-initiated actions", in *Cognitive contributions to the perception of spatial and temporal events*, eds. G. Aschersleben, T. Bachmann, and J. Müsseler (Amsterdam: Elsevier), 215-231.

Haggard, P., and Eimer, M. (1999). On the relation between brain potentials and awareness of voluntary movements. *Experimental Brain Research* 126,128–133.

Haggard, P., Newman, C., and Magno, E.(1999). On the perceived time of voluntary actions. *British Journal of Psychology* 90, 291–303.

Haikonen, P. O. (2003). *The Cognitive Approach to Conscious Machine*. Exeter, UK: Imprint Academic.

Harnad, S. (1990). The symbol grounding problem. *Physica D* 42, 335-346.

Hohwy, J. (2007). The Sense of self in the Phenomenology of Agency and Perception. *Psyche* 13, 1-20.

Iwasaki, S. (1993). Spatial attention and two modes of visual consciousness. *Cognition* 49, 211-233.

James, W. (1890). *The Principles of Psychology*. New York: Holt.

Johnson-Laird, P. N. (1988). "A computational analysis of consciousness", in *Consciousness in Contemporary Science*, eds. A. J. Marcel and E. Bisiach (Oxford. Oxford University Press) 357-368.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Karmiloff-Smith, A. (1992). *Beyond Modularity. A Developmental Perspecitive on Cognitive Science*. Cambridge, Massachusetts: MIT.

Knight, D. C., Nguyen, H. T., and Bandettini P. A.(2006). The role of awareness in delay and trace fear conditioning in humans. *Cognitive, Affective,and Behavioral Neuroscience* 6 (2), 157-162.

Knudsen, E. I. (2007). Fundamental Components of Attention. *Annual review of neuroscience* 30, 57-78.

Koch, C. and Tsuchiya, N. (2006). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences* 11 (1), 16-22.

Kouider, S. and Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Phil. Trans. R. Soc. B* 362, 857-875.

Kouider, S., de Gardelle, V., Sackur, J., Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences* 14 (7), 301-307.

Kunde, W., Kiesel, A. and Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition,* 88, 223-242.

La Berge, D. (1995). *Attentional Processing. The Brain's Art of Mindfulness*. MA. Harvard University Press.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences* 7, 12-18.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences* 10, 494-501.

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscince* 1(3), 204-240.

Large, E. W., Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychological Review* 106 (1), 119–159.

Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology* 11, 475-480.

Laughlin, S. B. and Sejnowski, T. J. (2003). Communication in Neuronal Networks. *Science* 301, 1870-1874.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance* 21, 451-468.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioral and Brain Sciences* 8, 529–566.

Libet, B. (2004). *Mind Time. The Temporal Factor in Consciousness.* Cambridge, Massachusetts: Harvard University Press.

Libet, B., Gleason, C. A., Wright, E. W., Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness potential): the unconscious intention of a freely voluntary act. *Brain* 106, 623–642.

Mach, E. (1890). *Contributions to the analysis of the sensations*. La Salle: The Open Court Publishing Company.

Mack, A. and Rock, I. (1998). *Inattentional Blindness.* MA: MIT Press.

Madl, T., Baars, B. J., Franklin, S. (2011). The Timing of the Cognitive Cycle. *PLoS ONE* 6(4), 1-16.

Mandler, J. M. (2008). On the Birth and Growth of Concepts. *Philosophical Psychology* 21, 207-230.

Mandler, J. M. (2010). The spatial foundations of the conceptual system. *Language and Cognition* 2, 21-44.

Marchetti, G. 2001. A Theory of Consciousness. Retrievable from: http://www.mind-consciousness-language.com/comm.htm.

Marchetti, G. 2005. Commentary on Benjamin Libet's *Mind Time. The Temporal Factor in Consciousness*. Retrievable from: http://www.mind-consciousness-language.com/comm.htm.

Marchetti, G. (2009). Studies on time: a proposal on how to get out of circularity. *Cognitive Processing* 10 (1), 7-40.

Marchetti, G. 2010. *Consciousness, Attention and Meaning*. New York: Nova Science Publishers.

Merikle, P. M. and Daneman, M. (1998). Psychological investigations of unconscious perception. *Journal of Consciousness Studies* 5, 5-18.

Merikle, P. M., Smilek, D, Eastwood, J. D (2001). Perception without awareness: perspectives from cognitive psychology. *Cognition* 79, 115-134.

Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology* 28, 597-613.

Metzinger, T. (2000). "The *Subjectivity* of Subjective Experience: A Representationalist Analysis of the First-Person Perspective", in *Neural Correlates of Consciousness. Empirical and Conceptual Questions*, ed. T. Metzinger (Cambridge. Massachusetts: The MIT Press) 285-306.

Nakayama, K. and Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research* 29, 1631-1647.

Negrotti, M. (1997). *La terza realtà. Introduzione alla teoria dell'artificiale.* Bari: Dedalo.

Negrotti, M. (1999). *The Theory of the Artificial. Virtual Replications and the Revenge of Reality*. Exeter, England: Intellect Books.

Newen, A. and Vogeley, K. (2003). Self-representation: Searching for a neural signature of self-consciousness. *Consciousness and Cognition* 12, 529-543.

Nisbett, R. E., and Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review* 84, 231–259.

Northoff, G. and Panksepp, J. (2008). The trans-species concept of self and the subcortical-cortical midline system. *TRENDS in cognitive sciences* Vol. 12, No. 7, 259-264.

Nummenmaa, L. and Calder, A. J. (2008). Neural mechanisms of social attention. *Trends in Cognitive Sciences* 13, 135-143.

Núñez, J. P. and de Vincente, F. (2004). Unconscious Learning. Conditioning to Subliminal Visual Stimuli. *The Spanish Journal of Psychology*, 7, 13-28.

Panksepp, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition* 14, 30-80.

Pashler, H. E. (1998). *The psychology of attention*. MA: MIT.

Pelphrey, K. A., Morris, J. P., and McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience* 16, 1706-1716.

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., and McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex* 15, 1866-1876.

Perruchet, P. (2008). "Implicit learning", in *Cognitive psychology of memory. Vol.2 of Learning and memory: A comprehensive reference*, ed. J. Byrne (Oxford: Elsevier) 597-621.

Perruchet, P. and Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences* 25(3), 297-388.

Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant.* Paris et Neuchâtel: Delachaux et Niestlé.

Piaget, J. (1937). *La construction du réel chez l'enfant.* Paris et Neuchâtel: Delachaux et Niestlé.

Piaget, J. (1974). *La prise de conscience.* Paris: Presses Universitaires de France.

Posner, M. I. (1990). "Hierarchical distributed networks in the neuropsychology of selective attention", in *Cognitive Neuropsychology and Neurolinguistics*, ed. A. Caramazza (Hillsdale, NJ.: Lawrence Erlbaum) 187-210.

Posner, M. I. (1994). Attention: The Mechanisms of Consciousness. *Proceedings of the National Academy of Sciences of the United States of America* 91, 7398-7403.

Posner, M. I. (1995). "Attention in cognitive neuroscience: an overview", in *Handbook of cognitive neuroscience*, ed. M. S. Gazzaniga (MA: MIT) 615-624.

Posner, M. I. and Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience* 13, 25-42.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* 6, 855–863.

Rees, G., Wojciulik, E., Clarke, K., Husain, M., Frith, C., and Driver, J. (2002). Neural Correlates of Conscious and Unconscious Vision in Parietal Extinction. *Neurocase* 8, 387-393.

Rey, A., Goldstein, R M., Perruchet, P. (2009) Does unconscious thought improve complex decision making? *Psychological Research* 73, 372-379.

Rizzolatti, G. and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience* 11, 264-274.

Rochat, P. (2003). Five levels of self-awareness as they unfold early in life. *Consciousness and Cognition* 12, 717-731.

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia* 46, 829-840.

Sackur, J. and Dehaene, S. (2009). The cognitive architecture for chaining of two mental operations. *Cognition* 111, 187-211.

Saxe, R. and Kanwisher, N. (2003). People thinking about thinking people: fMRI investigations of theory of mind. *NeuroImage* 19, 1835-1842.

Saxe, R. and Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia* 43, 1391-1399.

Scheider, S., Kuhn, W. (2011). Finite relativist geometry grounded in perceptual operations. *Conference on Spatial Information Theory 2011*.

Schibach, L., Eickhoff, S. B., Rotarska-Jagiela, A., Finkj, G. R., Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. *Consciousness and Cognition* 17, 457-467.

Searle, J. R. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3, 417-424.

Searle, J. R. (1984). *Minds, Brains, and Science: The 1984 Reith Lectures.* Cambridge: Harvard University Press.

Searle, J. R. (1992). *The Rediscovery of the Mind.* Cambridge: The MIT Press.

Seth, A. K. (2009). "Functions of consciousness", in *Encyclopedia of Consciousness, Vol. 1,* ed. W. P. Banks (Amsterdam: Elsevier) 279-293.

Shallice, T. (1988). "Information-processing models of consciousness: possibilities and problems", in *Consciousness in Contemporary Science*, eds. A. J. Marcel and E. Bisiach (Oxford. Oxford University Press) 305-333.

Shanks, D.R. (2005). "Implicit learning", in *Handbook of cognition*, eds. K. Lamberts and R. Goldstone (London: Sage) 202–220.

Shulman, R. G., Hyder, F., Rothman, D. L. (2009a). Baseline brain energy supports the state of consciousness. *PNAS* 106 (27): 11096-11101.

Shulman, R. G., Hyder, F., Rothman, D. L. (2009b). Brain Energy Supports the State of Consciousness. *Psyche* 15, 60-82.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Sciences* 14, 29-56.

Srinivasan, N. (2008). Interdependence of attention and consciousness. *Progress in Brain Research*, eds. R. Banerjee and B.K. Chakrabarti, 168, 65-75.

Stern, L.W (1897) Psychische Präsenzzeit. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* 13, 325–349.

Taylor, J. G. (2002). Paying attention to consciousness. *Trends in Cognitive Sciences* 6 (5), 206-210.

Taylor, J. G. (2007a). CODAM: A neural network model of consciousness. *Neural Networks* 20, 983-992.

Taylor, J. G. (2007b). "Through Machine Attention to Machine Consciousness", in *Artificial Consciousness*, eds. A. Chella and R. Manzotti (Exeter, UK: Imprint Academic) 24-47.

Thompson, E. (2007). Representationalism and phenomenology of mental imagery. *Synthese* 160, 397-415.

Tononi, G. (2007). "The information integration theory", in *The Blackwell Companion to Consciousness*, eds. M. Velmans and S. Schneider (Oxford, UK: Blackwell) 287-299.

Tsuchiya, N. and Adolphs, R. (2007). Emotion and Consciousness. *TRENDS in cognitive sciences* Vol. 11, No. 4, 158-167.

Ukhtomsky A. A. (1966). *The dominant*. Moscow–Leningrad: Nauka.

Umiltà, C. (1988). "The control operations of consciousness", in *Consciousness in Contemporary Science*, eds. A. J. Marcel and E. Bisiach (Oxford. Oxford University Press) 334-356.

Valéry, P. (1973a). *Cahiers*. Ed. by Judith Robinson-Valéry. Paris: Gallimard. (Italian translation, 1988: *Quaderni. Volume terzo*. Milano: Adelphi).

Vandekerckhove, M and Panksepp, J. (2009). The flow of anoetic to noetic and autonoetic consciousness: A vision of unknowing (anoetic) and knowing (noetic) consciousness in the remembrance of things past and imagined futures. *Consciousness and Cognition 18*, 1018-1028.

VanRullen, R., Carlson, T., Cavanagh, P. (2007). The blinking spot light of attention. *PNAS* 104, 19204-19209.

Varela, F.J. (1996). Neurophenomenology. *Journal of Consciousness Studies* 3, 330-349.

Varela, F.J. (1999). "The Specious Present: A Neurophenomenology of Time Consciousness", in *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, eds. J. Petitot, F. Varela, B. Pachoud, J.-M. Roy, (Stanford, CA. Stanford University Press) 266-329.

Varela, F.J. and Shear, J. (1999). "First-person Methodologies: What, Why, How?", in *The View from Within. First-person Approaches to the Study of Consciousness*, eds. F. J. Varela and J. Shear J. (Exeter, England: Imprint Academic) 1-14.

Varela, F. J., Toro, A., John, E. R., Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia* 19, 675–686.

Vicario, G. B. (1963). La "dislocazione temporale" nella percezione di successioni di stimoli discreti. *Rivista di Psicologia* 57: 17-87.

Vicario, G. B. (2005). *Il tempo. Saggio di psicologia sperimentale*. Bologna: Il Mulino.

Vogeley, K., Bussfeld, P., Newen, A., Hermann, S., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., and Zilles, K. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage* 14, 170-181.

Vogeley, K.,May, M., Ritzl, A., Falkai, P., Zilles, K., and Fink, G. R. (2004). Neural correlates of first-person-perspective as one constituent of human self-consciousness. *Journal of Cognitive Neuroscience* 16, 817-827.

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences* 7 (12), 553–559.

Waroquier, L., Marchiori, D., Klein, O., and Cleeremans, A. (2003). Is It Better to Think Unconsciously or to Trust Your First Impression? A

Reassessment of Unconscious Thought Theory. *Social Psychological and Personalty Science* 1 (2), 111–118.

Wickens, C. D. (1984). "Processing resources in attention", in *Varieties of attention*, eds. R. Parasuraman and D. R. Davis (New York: Academic Press) 63-102.

Zhong, C., Dijksterhuis, A., Galinsky, A. D (2008). The Merits of Unconscious Thought in Creativity. *Psychological Science* 19, 912-918.

In: Philosophy of Mind ISBN: 978-1-62257-215-1
Editors: R. J. Jenkins, W. E. Sullivan

Chapter 3

AN ANALYSIS OF THE NEUROPHYSIOLOGICAL EFFECTS OF QIGONG ON THE MIND

Yvonne W. Y. Chow
Department of Applied Social Sciences
The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

With the increasing demand of non-pharmacologic treatment for psychological problems, qigong has been shown to provide promising physiological and psychological effects. In view of the inadequate documentation analyzing the neurophysiological effects of qigong, this paper aimed at analyzing the underlying neurophysiological mechanism of qigong on the mind, especially on the brain activity. The brain activity under qigong state is quite different from that under other relaxed states (e.g. sleep and close-eyed rest), however, it is rarely mentioned in previous literature. To further understand this, studies employing neuroelectrical (e.g. EEG) and neuroimaging measurements (e.g. fMRI) were extracted from seven databases for analysis. Both Chinese and English written articles were included.

Findings showed that mind regulation and breath regulation of qigong training help to stabilize the autonomic and stress response systems. The unique neurophysiological mechanism of qigong, which is different from other relaxing mindful state, is characterized by activation of the parasympathetic system, improvement in cortical-subcortical

synchronization, thalamic performance, etc. We suggest that qigong can be a psychosomatic exercise of moderate intensity to stabilize the mind. However, due to the special nature of EEG and fMRI measurements, within-group design and relatively small sample size were commonly used in available studies. Therefore, more sophisticated RCT design with larger sample in future research is recommended for further justification.

Keywords: anxiety, body-mind, EEG, neurophysiological, qigong, stress

INTRODUCTION

Stress Related Illnesses

A considerable percentage of primary care doctor visits is related to stress (Mackenzie and Rakel, 2006). Stress can affect both the mind and the body. Acute stress stimulates the regulative reaction of neural, cardiovascular, autonomic, immune, and metabolic systems. However, when the body fails to cope with prolonged stressful conditions, dysfunctional consequences of these systems might be the result. Research evidence has shown, in modern societies, psychological stress is related to physical illnesses such as hypertension, heart disease, and cancer (Baum and Posluszny, 1999; Byrne and Espnes, 2008; Pandya, 1998; Stahl and Hauger, 1994; Tennant, 2000). Poor stress management is also linked to unhealthy behaviours, such as eating high-fat foods, not exercising, alcoholism, and smoking (McEwen, 2008; Ng and Jeffery, 2003). Psychological consequences, such as neuroticism, tension, depression, irritation, and anxiety have also been cited in previous literature (Beehr, 1976; Beehr, Walsh, and Taber, 1976; Caplan, Cobb, French, Jr. Van Harrison, and Pinneau, 1975; Coburn, 1975; Gemmil and Heisler, 1972). It is estimated that mood disorders affect approximately 20.9 million American adults, or about 9.5 percent of the U.S. population. The World Health Organization (WHO) (2001) estimated that mood disorders such as depression, will be the second leading cause of disabilities by the year 2020. It is estimated that 18% of American adults had some form of anxiety disorders (Kessler, Chiu, Demler, and Walters 2005). Moreover, anxiety has affected between 40% and 60% of the individuals who have heart conditions (De Jong, Moser, An, Chung, 2004; Denollet and Brutsaert, 1998; Konstam, Moser, and De Jong, 2005). A large proportion of disability, absentee from work and huge medical expenses increased the economic burden in different parts of the

world. The UK Confederation of British Industry (1996) estimated that 24 million working days were lost because of stress-related problems and it constituted about £1.3 billion of economic loss (Rose and Taylor, 1998). It is stated in Healthy People 2000, a report from the U.S. Department of Health and Human Services (2001), that two-thirds of visits to doctors are for stress-related illnesses and stress contributes to 50% of all illnesses in United States. About 65 million stress sufferers spend more than US $370 million a year on stress-related medications and treatments (Segail, 2007).

However, people commonly present their psychological problems as somatic discomforts, which include headache, lower back pain, neck pain, coughing, bronchitis, breathing difficulties, stomach pain, and abdominal problems (Brosschot, Gerin, and Thayer, 2006; Rose and Taylor, 1998). Somatization makes the general practitioners less likely to diagnose anxiety (Kessler, Lloyd, and Lewis, 1999). Less than half of the anxiety cases are identified in the primary care consultations (Goldberg and Huxley, 1992). Moreover, due to concerns about discrimination and loss of face in families (Acuna and Bolis, 2005), persons with mental illness often conceal their problems and do not seek the appropriate help at an early enough stage (Lee, Lee, Chiu, and Kleinman, 2005). Failure to seek prompt treatment might further exacerbate the problems (e.g. suicide) (Census and Statistics Dept., HK, 2008; Leung et al., 2009). Cognitive behavioural therapy (CBT) and drug therapy are two mainstream treatments of mood disorders. However, some studies showed that CBT is not always effective with patients suffering from anxiety disorders, especially for those who prefer medication (Grevier, 2003). Drug therapy is not always desirable as it carries side-effects (e.g., nausea), which usually brings a reduction in the quality of life (Croog, et al., 1986). It also requires good drug compliance and constant monitoring by physicians to enhance its effectiveness. Haslam and his colleagues (2003) found that, lack of improvement in symptoms and concerns about dependency made sufferers of anxiety and depression prematurely ceased medication. People are now more aware of the importance of maintain a health regime to prevent psychological and physical illnesses. Inadequacy of mainstream treatments further stimulate people to seek for self-help and easy to use body-mind strategies to promote body-mind welling. While it is widely known that regular aerobic exercises are beneficial to health, high intensity exercise might discourage the elderly population (Lan, Chou, Chen, Lai, and Wong, 2004), or those who dislike the running and sweating of regular exercise (Berger and Owen, 1988). It stimulates the involvement of mindful exercise (e.g., qigong or yoga) as a non-pharmacologic, home-based and less stigmatizing intervention to stabilize

mood disturbance at an early stage or as a complementary therapy to conventional treatment (Fang and Schinke, 2007; Lee, Lim, and Lee, 2004).

A Way Out - Mindful Exercise

Mindful exercise is different from most physical activities that centre on kinesthetic training. Mindful exercise is of low-to-moderate intensity focusing on self-awareness and intrapersonal mind-body alignment. Forge (2005) suggests that mindful exercise should include introspective meditation, proprioceptive awareness of bodily movement, breath-centering techniques, anatomic alignment postures, and contemplation of energy-centric flow. Yoga and qigong are examples of mindful exercise that maintain a belief in the "life energy," such as *prana* or *qi*, focus on inner mental activity and self awareness, synchronize muscular movements with breathing patterns, and pay attention to form and alignment (Forge, 1997). In fact, researchers suggested that a moderate exercise regimen can improve heart functions (Durham, 2005), and bring a mood elevation similar to that observed in athletes doing vigorous exercise (Berger and Owen, 1992). Netz and Lidor (2003) compared the effectiveness of mindful versus aerobic exercise modes in mood alternations. Their results indicate that low-intensity, cognitive focus, rhythmic movements may be more effective in immediate mood improvement than high-exertion kinesthetic-based movements.

It is commonly agreed that physical exercise has a positive impact on body-mind well-being. However, public awareness and the level of research attention focused on mindful exercise are much less than with many aerobic exercises (Daley, 2002; Scully, Kremer, Meade, Graham, and Dudgeon, 1998). Research evidence has shown that the practice of mindful exercises was related to improvement in psychological (e.g. stress, anxiety and depression) (Brown and Gerbarg, 2005; West, Otte, Geher, Johnson, and Mohr, 2004) and physical health (e.g. hypertension and chronic heart problem) (Bernardi, et al., 2002; Henderson, Hart, Lai, and Hunyor, 1998; Schell, Allolio, and Schonecke, 1994; Schneider, et al., 1995; Sundar, et al., 1984). Inhibition of the sympathetic nervous system has been shown to strengthen the immune system in several forms of meditation: including mindfulness meditation, qigong, and transcendental meditation (Collins and Dunn, 2005; Davidson, 2003; Lee, Huh, et al., 2003; Takahashi, et al., 2005). Some research efforts on elderly depression have demonstrated that qigong as a mindful exercise is a similarly promising alternative therapy for alleviating mood disorders through

improving the psychosocial health of the sufferers (Tsang, Cheung, Lak, 2002; Tsang, Fung, Chan, Lee, and Chan, 2006; Tsang, Mok, Au Yeung, and Chan, 2003). A randomized controlled trial indicated that improvement in self-regulation, concentration, social skills and control of tantrums were found in autistic children after qigong intervention (Silva, Cignolini, Warren, Budden and Gooch, 2007) The researchers suggested that qigong may trigger the self-soothing response, and allow the child's nervous system to learn to self-soothe. However, the neurological mechanism behind was not thoroughly discussed.

Qigong and Its Significance

Qigong is an ancient psychosomatic art of self-healing with a history of about five thousand years in Chinese culture (Yao, 1991; Zhang and Rose, 2001). "*Qi*" means vital energy and "*gong*" means discipline. The concept of *qi* includes the air we breathe in, *yuan qi* (innate vital substance), *jing qi* (*qi* that flows in the meridians), and the *qi* derived from nature. Besides practice, it is also regarded "*gong*" as the improved functions of internal organs through qigong practice (Zhang, 2005). Scientists have debated about the existence of *qi* over the last two decades. However, researchers have provided some evidence of *qi* by using physical signal detectors such as the far-infrared detectors, which detected a modified far-infrared radiation from the palm of a qigong practitioner (Chen, 2004). In Traditional Chinese Medicine (TCM), a balanced state of internal *qi* flow is essential for optimal physical and emotional health as well as vitality. If *qi* does not move and behave properly inside the body, the person will get sick. Qigong is a mindful exercise focusing on the balance of *yin* and *yang* , as well as enhancing the circulation of *qi* inside the body. Qigong is mainly composed of five elements: visualization, meditation, relaxation, deep breathing, and target *qi* circulation (Tseng, Lin, and Yeh, 1995). Qigong puts emphasis on the control of the mind-states, posture, and breath. Three self-disciplinary regulations (i.e. mind, body and breath) are essential. Diaphragmic breathing and *qi* mobilization are the focus of "breath regulation" (*tiao xi*). The "body regulation" (*tiao shen*) requires the participant to do alignment postures in a relaxed manner. "Mind regulation" (*tiao xin*) is key to achieving a *ru jing* (tranquility) state where thoughts should be *yi shou* (focused) and in harmony. Qigong is a Chinese exercise for *yang sheng* (health preservation) and *bao jian* (health maintenance). It shares the common health concept of harmony of *yin* and *yang* with Traditional Chinese

Medicine (TCM), which is more popular in the Chinese culture. It is suggested that qigong exercise is not costly; it is easy for everyone to practice. It helps people to prevent illness, preserve health, and enhance longevity (Li, 2009).

One of the prime benefits of qigong is stress reduction. The main catalyst which calms the brain is the qigong practitioner consciously focusing on a point (e.g. a part of the body) while maintaining a "let go" attitude. Previous qigong studies have documented the relaxation response produced by qigong practice and its effectiveness in reducing stress. After a period of qigong intervention, the physical changes can be observed in respiration (Lim Boone, Flarity, and Thompson, 1993), brain activity (Lee, et al., 1997), blood pressure, skin conductivity (Kuo, Ho and Lin, 2003), and heart rate variability (Yu, Shen and Chai, 2000). He, Le, Xi, and Zhang (1999) used a "stress meter" to measure the skin conductivity of the subjects and found that the degree of relaxation achieved by qigong meditation was significantly deeper than that of non-qigong meditation. A significant negative correlation was found between the length of qigong training and the subscales of Symptoms of Stress (SOS) (Lee, Ryu, and Chung, 2000). Studies that recruited patients as subjects indicated that qigong therapy was useful in treating anxiety (Shan, Yan, Sheng, and Hu, 1989) and depression in psychiatric settings (Pavek, 1988; Wang, 1993). It is also reported that qigong exercise could significantly reduce stress, anxiety, adrenocorticotropic hormone (ACTH), cortisol (a kind of **s**tress hormones) and aldosterone levels (Chow, Dorcas and Siu, 2012; Lee, Kang, Lim, and Lee, 2004). Improvement in mood, cellular function of neutrophil and natural killer cells were found as well (Lee, Hui, et al., 2001). These clinical reports support the scientific position of qigong as an effective mindful exercise to improve symptoms of anxiety and stress. It is suggested that qigong distinguishes itself from other therapies in terms of its TCM concept and emphasis of *qi* (vital energy) and *yi* (intention power), which manages our mood states both physically and psychologically, and achieves a result that goes beyond the simple reduction of stress (Lehrer, Woolfolk, and Sime, 2007).

Neurophysiological Effects of Qigong

However, we know little about the neurophysiological impacts of qigong on the brain. . Although neurophysiological studies of the brain activity of practitioners of qigong exercises have been conducted for about fifty years, the underlying neurophysiological effects of qigong have not been concluded yet.

So far, there is no clear consensus of how changes happen inside the brain (e.g. brain wave, neurotransmitters, etc.) while qigong practice calms the soul and body. Therefore, this paper aimed at analyzing the underlying neurophysiological mechanism of qigong on the mind, especially on the neural activity and brain functioning. Yoga and qigong are commonly regarded as two traditional origins of mindful exercises. However, the attention given to qigong is far more less than the research effort on yoga. The brain activity under qigong state is quite different from that under other relaxed states (e.g. sleep, rest, etc.). However, it is rarely documented in previous literature. To further understand this, in our review, studies focused on qigong and employed neuroelectrical (e.g. electroencephalography), neuroimaging measurements (e.g. functional magnetic resonance imaging) and neurochemical (e.g. blood test of hormones) were extracted from seven databases for analysis.

Reviews and comprehensive analysis are not common in the field of qigong studies. Qigong is rooted in Chinese culture. Early qigong studies were mostly done in China and were written in Chinese which might explain why comprehensive analysis is rare so far, especially in English written journals. Among the few reviews, Cahn and Polich (2006) reviewed the neurophysiological findings of different mindfulness-based practices such as yoga, hypnosis and qigong. They concluded that these exercises demonstrated promising clinical effects on reducing anxiety and other stress -related disorders. Their review also provided a frame-work for qualitative analysis of qigong effects. However, only four qigong studies were included in that review which was hardly enough to conclude the neurophysiological effects of the qigong mindfulness state. Ng and Tsang (2009) reviewed 26 randomized controlled trials of qigong. This meta-analytical review provided some positive evidence of the benefits of health qigong which made qigong less regarded as "pseudoscience" as before. However, their review only covered studies in a recent ten year period (i.e. 1997-2009). The physiological measurements mainly focused on blood and urine tests, neuroelectrical parameters (e.g. brain wave) were not included in their review. The earliest studies focusing on EEG performance of qigong participants were done in China about 50 years ago (Qin and Zhai, 1962). Therefore, quite a lot of valuable information might have been missed if those earlier Chinese studies are not included. Guo and his colleagues (2008) have done a meta-analysis on nine randomized clinical trials of qigong. This review provided more evidence that qigong combined with drugs has better effect than drug therapy alone, especially on reduction of hypertension symptoms and improvement of quality of life. However, the

focus was restricted to the clinical effect on essential hypertension as reflected in blood pressure, heart rate and respiration. Thus, the neurological aspect was not highlighted. The human body is a highly complex system of energies including fluctuating mood states as well as the nervous, lymphatic, motor and endocrine systems. They all play a vital and cooperative part in the body-mind wellbeing of a person, and maintaining a state of homeostasis. Previous reviews are rather sketchy with limited coverage which may overlook the contribution of earlier studies. Therefore, it is still difficult to provide an overall picture of the qigong effects from a holistic body-mind perspective. Moreover, in view of relatively less attention given to the neurophysiological effects of qigong on mental health, we therefore, aim at reviewing more neurological studies of qigong in this preliminary review and to synthesizing the underlying psycho-neurophysiologic effects of qigong on stress regulation.

METHODS

Data Sources and Search Strategies

In the studies that investigated the neurophysiological effects of qigong were often measured using instruments under four main categories: i.e. neuroimaging, neuroelectrical, neurochemical and electrodermal instruments. Neuroelectric studies included methods such as electrocephalography (EEG) and event-related potential (ERP) which detect the brain's electrical activity. Neuroimaging measurements such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) record the increase in blood flow to the local vasculature that accompanies neural activity in the brain (Cahn and Polich, 2006; Kokubo, 2001). Neurochemical techniques collect blood samples and test the concentration of hormones, endocrine and neurotransmitters which reflect the autonomic stress response (Lee, Huh, et al., 2001; Lee, Kang, et al., 2004). Electrodermal assessments (e.g. high-voltage electrophotographic camera) capture the electromagnetic radiation, wavelength and luminous phenomena generated from the body. Questionnaires were sometimes included to measure the psychological aspects (e.g. perceived stress). We searched for articles indicating to investigate qigong effects by direct measures of neural activity and brain functioning. Neuroimaging, neuroelectrical, neurochemical and electrodermal were used as the initial keywords as studies with these words indicate that neurophysiological measures were likely conducted. The list of keywords were

further expanded to include *qi*, qigong (氣功 for Chinese database), chi-gong, *qi*-training, qigong state, EEG, ERP, fMRI, PET, electromagnetic, electromyographic, immunological, neuroendocrine, neurohormonal, neurotransmitters, anxiety, stress, cortisol, noradrenaline, ACTH, brain wave, cognitive, sensory, etc. We searched Medline, PsyINFO, PsyARTICLES, ProQuest, and Science Direct (Elsevier) for relevant articles from 1950 to 2010. Such an extended period might include those Chinese pioneers in the field which are beyond the inclusion scope of the review by Ng and Tsang (2009). Though a number of studies have been conducted in China and the earliest can be traced back to the 1960's, very few have been reviewed in western publications. One possible reason might be that those research articles are written in Chinese. Therefore, we also searched Chinese Social Sciences Citation Index and CAJ full-text database to identify the neurological articles written in Chinese.

Studies that used standardized measuring instruments, focused on the neurophysiological effects of qigong and its influence on stress and/or anxiety reduction and mood regulation were included in this review. Inclusion criteria included: (a) employing one experimental group and at least one control group for comparison, (b) using self-practiced qigong or qigong therapy as main intervention which could be either individual-based or group-based, (c) with outcome measures containing neurophysiological parameters (e.g. changes in brain wave, blood test, etc.) or brain functioning (sensory, cognitive processing, etc.), and/or (d) psychological conditions (e.g. Spielberger anxiety inventory, perceived stress scale, etc.). We excluded case studies and studies with small sample size ($n < 5$). Those experiments involving one time *qi*-emission by qigong masters were not included in this review.

Data Analysis

We followed the method used in Cahn and Polich's article (2006) to analyze the neurological effects according to the nature of impact (e.g. brain waves, neurotransimitters, etc.) which were measured by different means (e.g. neuroimaging, neurochemical, etc.). We also compared the brain wave activities between the qigong state, a remarkable state with special excitation brain activities (Zhang, Li and He, 1988; Zhang, Zhao and He, 1988) and other relaxation states (i.e. open-eyed rest and sleep).

RESULTS

Initial keyword search yielded 115 articles. Fifty- one articles were excluded in the first screening stage as they focused solely on physical illnesses (e.g. cancer) without discussing the impact of qigong on stress alleviation and mood regulation. Twenty-seven articles were further excluded because comparison group was not available (e.g. within-group design using before and after comparison). Ten studies did not fulfill our inclusion criteria as they were either case studies or with small sample size. Four studies used one time *qi*-emission by qigong master and five non-intervention publications (e.g. review articles) were also excluded. We excluded one more study because of the inadequacy of data and absence of discussion in the article. Finally a total of seventeen research articles which satisfied our inclusion criteria were identified. Table 1 shows the details of these 17 selected studies. All 17 studies used control groups for between group comparisons. Five of them are randomized clinical trials (RCT) while two mentioned single-blinded setting (Chan, Cheung, Sze, Leung and Shi,. 2011; Lee, Huh, et al., 2001; Lee, Lee, Kim and Moon, 2003; Silva, et al., 2007; Yu, et al., 2005) Eight studies employed both within-group and between-group comparisons (Chan, et al., 2011; Kawano, Koita, Fujiki and Shinagawa, 1990; Lee, Huh, et al., 2001; Qin , Jin and Lin, 2009; Shen, et al.,2008; Sun, Lei, Yan and He, 1989; Zhang Chen and Xu, 2008; Zhang, Li, et al., 1988).

Nine neuroelectric studies used EEG to signal the changes in brain wave (Chan, et al., 2011; Fong, Hung and Huang, 2004; Kawano,et al.,1990; Pan, Zhang and Xian, 1994; Qin, et al., 2009; Shen, et al., 2008; Sun, et al., 1984; Sun, et al., 1989; Zhang, Li, et al., 1988) while one used surface electromyography (sEMG) to detect the muscle action potentials from underlying skeletal muscles (Wirht, Cram, and Chang, 1997). One neuroimaging study used brain electrical activity mapping (BEAM) (Zhang, et al., 2008) and one used fMRI technology (Yu, et al., 2005) to see images of changing blood flow in the brain associated with neural activity. Four studies employed neurochemical techniques (e.g. urine or blood test) to detect the changes in hormones and neurotransmitters (Lee, Lee, et al., 2003; Lee, Huh, et al., 2001; Lee, Kang, et al., 2004; Skoglund and Jansson, 2007). The researchers in five studies also used questionnaires to assess the psychological conditions (e.g. stress level) of the subjects (Lee, Huh, et al., 2001; Lee, Kang, et al., 2004; Lee, Lee, et al., 2003; Silva, et al., 2007; Skoglund and Jansson, 2007). Details of the 17 studies are illustrated in Table 1.

Regarding the neuroelectrical changes in brain activity during qigong practice, six studies found there were increased frontal alpha activity with larger amplitude and lower frequency under qigong state (Fong, et al., 2004; Kawano, et al., 1990; Qin, et al., 2009; Sun, et al., 1984; Yu, et al., 2005; Zhang, Li, et al., 1988). One study found decreased alpha activity in the whole brain, especially in the occipital region (Zhang, et al., 2008). Three studies mentioned better brain synchronization, especially between frontal and occipital zones (Kawano, et al.; Sun, et al., 1984; Yu, et al., 2005). Moreover, one study mentioned bilateral brain synchronization (Qin, et al., 2009) and one study found enhanced temporal alpha asymmetry after a short period of qigong practice (Chan, et al., 2011). Three studies found increased theta rhythm in frontal channels during qigong practice (Fong, et al., 2004; Pan, et al., 1994; Qin, et al., 2009). One neuroelectric study found qigong subjects were more able to concentrate and were less distracted than non-qigong subjects (Kawano, et al.).

The results of one study indicated a significant rise in neuromuscular activity during qigong therapy (Wirht, et al., 1997). Furthermore, the neurochemical findings in four studies indicated improved regulation of neurotransmitters and stress hormones, reduced anxiety and stress levels. (Lee, Huh, et al., 2001; Lee, Kang, et al., 2004; Lee, Lee, et al., 2003; Skoglund and Jansson, 2007). Other characteristics of enhanced parasympathetic dominance such as decreased heart rate and finger temperature were obtained in one study (Skoglund and Jansson).

There were six longitudinal studies that required the subjects to practice qigong for a period ranging from four weeks to six months (Chan, et al., 2011; Lee, Lee, et al., 2003; Lee, Kang, et al., 2004; Silva, et al., 2007; Skoglund and Jansson, 2007; Sun, et al., 1989). Qigong exercise yielded promising effects on easing stress related symptoms (e.g. cortisol reduction) (Lee, Kang, et al., 2004; Lee, Lee, et al., 2003; Skoglund and Jansson) One study demonstrated improved sensory functioning (Silva, et al.) and one found increased fluid intelligence (Sun, et al., 1989) after qigong intervention. Moreover, the uniqueness of brain activity in the qigong state was distinguished from the state of normal rest and sleep in three studies (Kawano, et al., 1990; Sun, et al., 1989; Yu, et al., 2005).

Table 1. Summary of 17 Studies (in chronological order)

	Study	Types of qigong (qg)	Aim	Participants N (M:F) (mean age; range)	Nature of Participants	Experimental design between gp (b); within gp (w); group (n) ; state (mins)	Tools	Findings
1	Sun, Wang, Liu, et al, (1984)	Neiyang gong	Analyze EEG power specturn and coherence during qigong	34 (26M:8F) 39 (29M:10F) 5 (4M:1F)	34 qi-trainees 39 non-qi trainees 5 non-qi trainees	b: qigong (34) vs rest (39) vs sleep (5) vs	EEG	-decreased of δ power - increased α power value - increased anterior and posterior coherence
2	Zhang, Li, He. (1988)	not specified	Show whether EEG changes related to time of qigong practice	10 (48.2y) 10 (56y) 10 (52y)	10 veterans (10-47yrs) 10 novices (2-3yrs) 10 control subj (never)	b: compare 3 gps w: pre(5mins) vs qigong (10mins) vs post (5mins)	EEG	- veterans, rest: α-2 predominant in posterior, α-1 even distributed; qigong: -α-2 decrease, α-1 increase in anterior; α rhythm moved fr posterior to frontal. - no sign change in novices and control gp
3	Sun, Lei, Yan, et al. (1989)	Zhanzhuang gong, meditative qg	Measure qigong effect on fluid intelligence and EEG physical age	C-:48(32M:16F) (62.65y) 48(36M:12F) (61.63y) L-T :19 (12M:7F) (66.63y)	48 qi-trainees 48 non-qi trainees 19 healthy persons	b: cross-sectional ; qi-gp (48) vs non-qi gp (48) w: b and a 6-mth qigong	EEG	- qigong gp were 1.62y younger than non-qigong gp - subjects became 1.32y younger after practiced qigong

	Study	Types of qigong (qg)	Aim	Participants N (M:F) (mean age; range)	Nature of Participants	Experimental design between gp (b); within gp (w); group (n) ; state (mins)	Tools	Findings
4	Kawano, Koita, Fujiki, et al. (1990)	Not specified	Analyze neurophysical effects of qigong	18(17M:2F) (43.5y; 21-73y) 19(16M:3F) (30.7y;19-50y)	18 qi- trainees 19 non-qi trainees	b: compare 2 gps w: rest vs qigong vs post	EEG	- rest: trainees show small α power and diffused toward frontal; non-trainees show large occipital α - Qigong: 2 gps show increased frontal α and resembled to occipital
5	Pan, Zhang, Xian. (1994)	concentrative , non-concentrative qg	Compare θ waves of concentrative qigong (CQ) and Non-CQ	20 (17M;3F) (29y;19-52y) 30 (22M;8F) (26y;16-46y) 23 (15M;8F) (28y;20-38y)	50 qi-trainees 23 non- qi trainees	b: CQ (20) vs NCQ (20) vs control (23)	EEG	- Fmθ related to mental activities in concentrative qigong practice

Table 1. (Continued)

	Study	Types of qigong (qg)	Aim	Participants N (M:F) (mean age; range)	Nature of Participants	Experimental design between gp (b); within gp (w); group (n) ; state (mins)	Tools	Findings
6	Wirth, Cram, Chang. (1997)	qigong therapy	Measure the variable energizing effects of qigong	17 14 14 (31.6y; 18-46y)	17 qi-trainees (patients) 14 TT-trainees (patients) 14 non-believers	b: compare 3 gps; @ gp rest + qigong/TT + rest + qigong/TT (total 20min)	sEMG	- qigong: rise in neuromuscular energy - TT: modest drop in overall muscle energy
7	Lee, Huh, Hong, et al. (2001)	ChunDoSunBup	Investigate psycho-neuroimmunological effects of qigong	10 (M) (26.7y) (qi-therapy) 10 (M) (25.4y) (placebo)	healthy non-qi trainees	b: random into qi- therapy (10min) vs placebo w: pre vs post (5 and 60 min)	Blood test, Spielberger anxiety Inv.	- melatonin increased in qi-therapy but not placebo gp - state anxiety decreased in qi-therapy gp
8	Lee, Lee, Kim, et al. (2003)	not specified	Investigate efficacy of qigong on hypertension	29 (56.0y) (qi-gp) 29 (56.5y) (control)	patients with hypertension	b: randomized into - qi-gp (29) vs control (29); b and a 10wk qigong practice	Blood test; Perceived stress scale	- significant reduction of NE, adrenaline, cortisol, blood pressure and stress level in qigong gp

	Study	Types of qigong (qg)	Aim	Participants N (M:F) (mean age; range)	Nature of Participants	Experimental design between gp (b); within gp (w); group (n) ; state (mins)	Tools	Findings
9	Fong, Hung, Huang. (2004)	meditative qigong	Investigate brain activity under qigong state and effects on emotion	15 15	healthy persons	b: qi-gp (15) vs control(15) b and a 20 min Qg/rest session	EEG	- qi-gp: α wave increased in frontal zone; power higher than control gp; β wave intensity less than control gp; θ wave power increased and higher than control gp
	Study	Types of qigong (qg)	Aim	Participants N (M:F) (mean age; range)	Nature of Participants	Experimental design between gp (b); within gp (w); group (n) ; state (mins)	Tools	Findings
10	Lee, Kang, Lim, et al.(2004)	ChunDoSunBup	Investigate effects of qi-training on anxiety, cortisol and ACTH	16(M) (30.5y;20-40y) (qi-)16(M) (31.2y;20-40y)(sham)	healthy persons	b: randomized into - qi-gp (16) vs sham control (16) b and a 4 wk qigong	Blood test; STAI-X1	- qi-gp: anxiety level decreased - ACTH, cortisol and aldosterone declined significantly
11	Yu, Li, Tang, et al (2005)	meditative qg	Compare low frequency synchrony btn qi- and rest; btn masters and control	5 (4M, 1F) 5 (4M,1F)	5 qigong masters 5 non-qi trainees	b: masters vs control in qigong and rest states	fMRI	- qigong state: low freq synchrony of the brain motor cortex was enhanced - increase α wave in frontal regions of qigong masters

Table 1. (Continued)

	Study	Types of qigong (qg)	Aim	Participants N (M:F) (mean age; range)	Nature of Participants	Experimental design between gp (b); within gp (w); group (n) ; state (mins)	Tools	Findings
12	Silva, Cignolini, Warren, et al. (2007)	Cignolini qigong protocol	Investigate whether qigong help sensory, cognition and social skills	15 (13M:2F) (3 – 6y)	kids with uncomplicated autism	b: randomized into qi-gp (8) and control (7) ; single blinded; b and a 5m qigong	questionnaire	- qi- gp: improved sensory impairment - increased social and basic living skills - improved bowel and sleep abnormalities
13	Shen, Shinlin, Zhu, et al. (2008)	San-tiao, meditative qg	Explore EEG changes at qigong meditation	8 11 (25-70y)	8 qi-trainees 11 non-qi trainees	b: qi-gp (8) vs control (11); b and a 1 Qg session w: rest vs qigong (15min)	EEG	- attentive and focus increased under qigong meditation in all 19 subjects; concentration of qi-trainees sustain longer and less distract than non-qi trainees (control)
14	Skoglund, Jansson (2007)	not specified	Investigate the effects of qigong on stress among computer operators	9 (F) (46y; 30-58y) 10 (F) (44y; 20-53y)	healthy computer operators	b: qigong (9) vs control (10); b and a 5 wk qigong	Electric cuff Urine test questionnaire	- qi-gp: NE excretion in urine reduced - heart rate and finger temp decreased - perceived stress reduced

	Study	Types of qigong (qg)	Aim	Participants N (M:F) (mean age; range)	Nature of Participants	Experimental design between gp (b); within gp (w); group (n) ; state (mins)	Tools	Findings
15	Zhang, Chen, Xu, (2008)	Neiyang gong	Investigate the brain activity by using BEAM	17 (45y) 5 (34y)	17 qi-trainees 5 non-qi trainees	b: qi-gp vs control w: b and a maths task and qigong meditate (30mins)	BEAM	- maths task: qi-gp increased frontal θ and δ; decreased central α; increased temporal and occipital β - qigong: decreased θ and δ in frontal; decreased occipital α and β
16	Qin, Jin, Lin. (2009)	meditative qigong	Investigate whether 45yrs qigong have long-term EEG changes	1 (M) 76yr 2 (M) 72 and76 yrs 6 (4M:2F)	45yr qi-veteran meditator controls non-meditator controls	b: qi-veteran vs controls w: pre-(rest) vs qigong (30 mins) vs post-	EEG	- qigong: veterans enhanced alpha in amplitude and duration; θ and β increased; bilateral synchronization and sharp contoured α; 8-Hz α spread from anterior fronto-central midline to entire brain
17	Chan, Cheung, Sze, et al., (2011)	Shaolin Dan Tian Breathing (DTB), both active and passive form	Investigate whether qigong would induce both relaxed and attentive states	22 (5M:17F) (qigong) 20 (5M:15F) (PMR) (25-64y)	healthy persons	b: randomized into qigong) vs PMR; b and a 4 wk Qg; single blinded w: pre-(rest) vs post passive Qg (5 mins) vs post active Qg (5 mins)	EEG	- qigong: enhanced temporal alpha asymmetry (i.e. relaxation and positive mood) after 5-mins passive DTB; enhanced intra- and inter-hemispheric theta coherence (i.e. attention and alertness) after 5-mins. active DTB

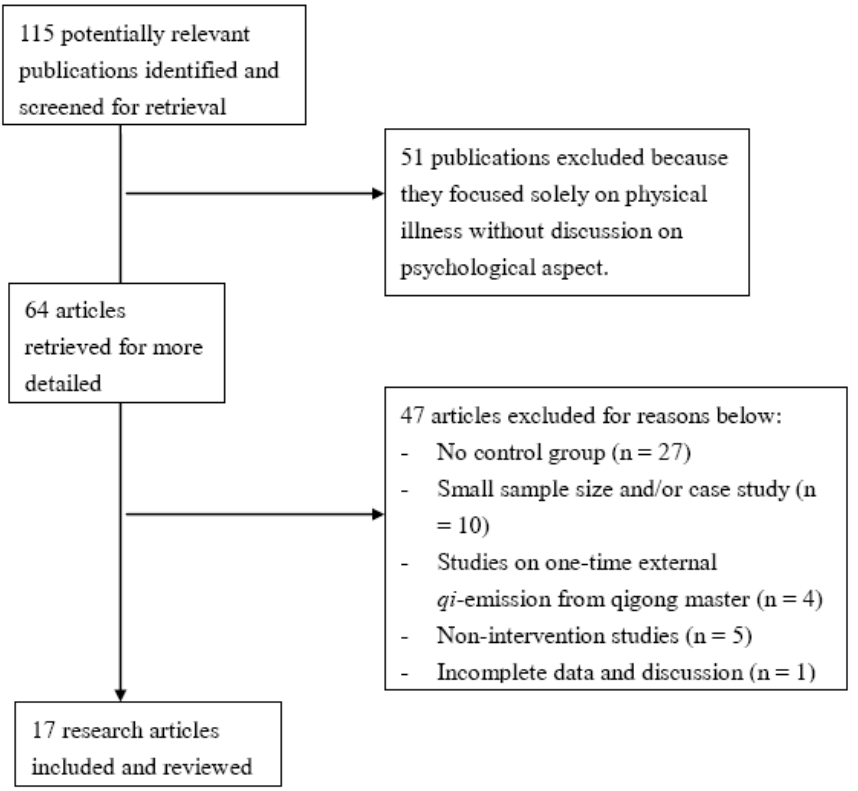


Figure 1. Flowchart of study selection process.

DISCUSSION

Neurological Effects of Qigong on Brain Activity

Qigong and Resting State

It is found, when at rest with eyes closed, the theta wave with lower power (e.g. 1.8 μv^2) mainly appears in the frontal area of the brain. In the qigong state, the duration of theta wave increased in the frontal cortex, with a slower frequency but higher amplitude (e.g. 2.7μv^2) (Qin, et al., 2009). The theta wave also gradually spread to other parts of the brain. A recent study found that, after 5 minutes of qigong practice, there was enhanced intra- and inter-hermispheric theta coherence The researchers suggested that such brain condition can be regarded as an indicator of attention and alertness state of

mind (Chan, et al., 2011). Thus, the qigong state is not simply resting but involves complicated interactions between cortical and subcortical regions. Regarding the alpha wave, it is also demonstrated that the frequency of alpha wave decreased from occipitally dominant 8.5-12Hz during the eye-closed rest state to frontally dominant 8-8.5Hz during qigong practice (Qin, et al., 2009; Zhang, Li, et al., 1988). Increased alpha rhythm possibly represented a more relaxed mental state with positive emotions induced by the "mind regulation" component and "let go" attitude of qigong practice (Fong, et al., 2004). The higher amplitude of low frequency fluctuation and increased dominant frequency in the brain cortex under qigong state as indicated in a fMRI study reflected that higher concentration is likely to be related to *yi shou* (Yu, et al., 2005). It is found that alpha frequency increasingly shifted to low frequency as the time of qigong state increased. The ultimate resonance of alpha wave between Fz and Oz represented an altered state of consciousness (ASC) where the alpha waves showed better synchronization from occipital to frontal zones (Kawano, et al. 1990; Yu, et al., 2005). The stronger the frontal alpha wave, the stronger the connections between the frontal lobe and the hypothalamo-pituitary section which leads to the improvement of brain performance (Qin, et al., 2009). It possibly help to reduce stress and anxiety levels, to enhance the parasympathetic tone and to regulate the psycho-neurophysiologic functioning. The BEAM study further found that alpha activity reduced in the whole brain, especially in the occipital region when under the states of *ru jing* and *yi shou* (Zhang, et al., 2008). The researchers suggested that *ru jing* is not a normal resting state but a special quiet state with cognitive activity. The *yi shou* practice involves selective attention (e.g *dantian*), sustained concentration, and emotional regulation. The distinctive theta and alpha activities in qigong state might indicate the experience of internal attention with positive emotions under the background of calm mental state. It is also suggested that such unique EEG profiles might ultimately become a biomarker of the qigong practitioner which requires further longitudinal justification (Qin, et al., 2009).

Qigong and Sleeping State

The initial stage of sleep is characterized by increased theta wave accompanied with gradually disappearing alpha waves. It is found when sleep proceeded to deeper stages, alpha wave decreased and delta wave increased in the occipital region (Sun, et al., 1984). However, in the qigong state, frontal midline theta waves increased under the background of increased amplitude and cycle duration of alpha wave (Fong, et al., 2004; Lee, Choi and Yook, 2005). It is suggested that increased anterior frontal and midline theta

synchronization was related to an attentive and blissful state (Chan, et al., 2011). It is also found that while alpha wave increased, delta wave decreased in the mid-frontal area (Sun, et al., 1984). Therefore the subjects in qigong state were under a relaxed psychological state with positive mood, but not in a sleeping, inattentive or drowsy condition. The coherence values at definite frequencies between frontal and occipital parts of the brain increased in the qigong state but decreased in the sleeping condition. The apparent changes in the frontal EEG spectrum indicated that there were mutually stimulating activities between cortical and subcortical brain areas. There are stimuli from internal organs mediated by the thalamus (the central distributor of information) and then transmitted to higher (cerebral) nerve centers. One possible explanation is that the state of *ru jing* quiets the brain and inhibits the cerebral cortex to interact with the external stimulation which ultimately releases the brain from executive functions (e.g. anticipation, planning and worry). The deepened brain activity further enhances the parasympathetic link from the hypothalamus to the autonomic nervous system in regulating the viscera. This mechanism stabilizes the anxiety arousal and contributes to the reduction of negative mood states (Lee, et al., 2002). It might explain how mental activities of the qigong state, which are different from the sleeping state, could influence the biological functioning of the body.

Concentrative Qigong (CQ) and Non-Concentrative Qigong (NCQ)

Different forms of qigong have different emphases on the three main disciplines of qigong. The CQ highlights the strategy of mind-regulation (i.e. *yi shou*) while the NCQ stresses on relaxation, positive mood and no-thought. Thus, mental activity in CQ is more active than in NCQ. Pan and his colleagues (1994) compared CQ and NCQ regarding the differences in frontal mid-line theta rhythm (Fmθ). The long rhythmicθtrains (i.e. frontal mid-lineθpeaks) were found frequently in CQ subjects but seldom in NCQ group. Distinct Fmθ is regarded as an indicator of mental activity and mental concentration in waking rest states and meditative states with positive emotions (e.g. blissfulness, relief) (Aftanas and Golocheikine. 2001). It is suggested that increased Fmθ was probably due to attentive mental activity in a rather relaxed CQ state. Scattered θ activity in NCQ practice may relate to the decreased arousal resulting from relaxation, peaceful mood state and no-thought (Pan, et al., 1994).

Implications from Neuroelectrical Findings

It is concluded that the qigong state was neither consciousness between wakefulness and resting drowsiness nor an unconscious sleeping state (Yu, et al., 2005). The changes in brain waves suggested there might be a more complicated interaction between cortical and subcortical regions. Such a unique state was characterized by a generalized inhibition of the cerebral cortex, increased alpha power with slower frequency in the anterior half of the brain and brain motor cortex. Decreased occipital alpha suggested a unique tranquility state with cognitive activity. Enhanced brain functional connectivity is characterized by the presence of frontal-occipital synchronization. This special activated state allowed the brain to process autonomic self- regulation, especially beneficial to sympathovagel augmentation and stabilization of emotional states.

Possible Bi-Directional Regulatory Effects of Qigong

Moreover, it is suggested that qigong has bi-directional regulatory potential (Zhang, Fu, and Zhao, 1990). On one hand, qigong helped to lower the NE blood concentration for those who suffered from essential hypertension (Lee, Lee, et al., 2003) and busy computer operators (Skoglund and Jansson, 2007) who ultimately demonstrated stress reduction. On the other hand, Liu and his colleagues (1990) found that qigong intervention helped patients who suffered from chronic pain and psychotic illness to increase levels of NE and dopamine (DA) resulting in improved mood, cognitive activities and positive psychological functioning. Their 5-HT concentration also reduced to normal level which helped in reducing sensitivity to physical painfulness, enabled adequate blood supply to the tissues and internal organs and thus facilitated the recovery speed. It seems rhythmical *dantian* breathing can improve ventilatory efficiency, mood stabilization, and flexibility of ANS via neuroendorcrine modulation (Lan, Chou, Chen, Lai, Wong, 2004; Lee, Lee, et al., 2003; Lim, et al., 1993). Qigong probably regulates and balances the neuroendorcrine response according to different physical and mental conditions of the individual, and finally achieves an equilibrium condition. In TCM, a balanced internal state is regarded as a healthy condition. However, no control group was used for more sophisticated comparison in Liu's (1990) study. Therefore, further investigation is required for the bi-directional regulatory potential of qigong.

Neurological Effect of Qigong on Hypothalamic-Pituitary – Adrenal (HPA) System

The emotional condition (e.g. anxiety level) might result from the complex interactions between neuroendocrine and immune systems through hormonal and neural channels. When one is stressed, an increase in heart rate and the release of stress hormones are amplified bodily reactions due to the activation of the sympathetic system (Lindforts and Lundberg, 2002; Lovallo, 2005; McKinney, Antoni, Kumar, Tims and McCabe, 1997). However, persistent and excessive amounts of stress hormones can weaken the immune system (Bloomfield, 1998; Lee, Huh, et al., 2001). After qigong intervention, reduced anxiety levels were reflected by decreases in blood pressure, noradrenaline, adrenaline, ACTH and cortisol levels (Lee, Lee, et al., 2003; Lee, Huh, et al., 2001; Lee, Kang, et al., 2004). The HPA axis is highly responsive for the inhibitory input of ACTH which consequently lowers the level of cortisol. Qigong exercise may modify the hypothalamic-pituitary-andrenocortical system for long-term stress coping. Skoglund and Jansson (2007) further found that qigong training may lead to a more efficient psycho-neurophysiologic network which was manifested by the reduction in heart rate, finger temperature and emotional irritability. Therefore, results from neurochemical analysis indicated that qigong possibly augments the stress response via neurohormonal modulation and thus achieves a so-called homeostasis condition (Lee, Kang, et al., 2004; Lee, Kim, et al., 2000). From a psycho-neurophysiologic perspective, it seems that qigong exercise is a psychosomatic training that enhances body-mind wellness via a complex psycho-neurological network which needs further scientific justification.

Neurological Effect of Qigong on Individual Psychological Traits

Personality

Evidence shows that personality traits as well as anxiety levels are related to Fmθ. Extroverts demonstrated a significantly higher amount of Fmθ than introverts. Extroverts also scored the lowest scores in neurotic traits and anxiety (Hashimoto, Mukasa, Yamada, Nakamura and Inanaga, 1988; Inanaga, 1998). The low Fmθ groups showed the opposite correlation. It is likely that the presence of Fmθ during qigong practice is related to a more outgoing personality, decreased pessimism and anxious emotions. The relaxing social environment and non-competitive characteristics of qigong might have a beneficial impact on enhancing social competence, extroversion and

friendliness (Lee, 2001; Leung and Singhal, 2004; Tang, 1995; Yuen, 2000). Moreover, if a person has emotional fluctuations, changes in electromyogram (EMG), finger pulse amplitude (FPA) and skin conductance levels (SCL) can be recorded. Optimistic suggestions and a non-striving attitude in the "mind regulating" discipline of qigong might have positive effects on stabilizing the emotional traits of the practitioners. However, an electromyographic study documented a significant rise in neuromuscular energy primarily concentrated in the L3 (lumbar) region of the patients during qigong therapy (Wirth, et al., 1997). The breath regulation of qigong always focuses on *qi* circulation inside the body so as to attain a balanced state (Takahashi and Brown, 1988). According to TCM, the lumbar region is regarded as a link to *dantian* (the center of energy) where *qi* is generated and accumulated. It is postulated that the focused *qi* practice stimulates the *qi* stored in *dantian* during treatment which gives rise to the EMG response. Therefore, the dynamic neurophysiological changes of qigong might express itself differently in EMG measures which warrants further investigation to explain such complex bodily mechanisms.

Sensory Processing and Intellectual Ability

It is suggested that as the power of alpha waves increase in the frontal region of the brain, the connections between the frontal lobe and the hypothalamo-pituitary section also improve which results in better brain functioning. Autistic children showed normalization of sensory impairments; improved performance in social skills and basic living skills after receiving qigong intervention (Silva, et al., 2007). The relaxing mental state of qigong practice possibly induces an increase in alpha rhythm which can be beneficial to the enhancement of cognitive abilities.

Poor fluid intelligence is a common result of declined brain functions in the elderly. It is found that qigong had positive effects on delaying the decline of fluid intelligence in the elderly (Sun, et al., 1989). The physical EEG age of qigong practitioners was 1.62 years younger than their own chronological age. They performed much better in arithmetic calculation, speed, short-term memory, and eye-hand coordination than the non-qigong group. Moreover, the physical age of EEG showed a reduction of 1.32 years after qigong intervention. Mastering difficult tasks in qigong practice possibly helps the elderly to improve self-efficacy. *Yi shou* requires a high concentration on breathing while clearing wandering thoughts. This might enhance the cognitive efficiency of the participants as well (Brekke, Hjortdahl and Kvien, 2003; Daley, 2002). However, when compared with the group using TCM

herbs to improve intelligence, the effect of qigong was 5 months slower. Therefore, further studies are needed to conclude whether qigong intervention needs longer time to achieve the same results and is more effective than other therapies.

CONCLUSION AND LIMITATIONS

Over stressed person is more vulnerable to physical and mental illnesses. The complex neurophysiological response involves the cortical and subcortical systems. Mechanisms leading to the stabilization of emotions include strengthened parasympathetic activity, stablized neurohormonal response, calm and smoothed cortical-subcortical coordination (i.e. synchronization), and activation of the limbic system which stimulates the forebrain reward reaction and emotional release (Brown and Gerbarg, 2005). Our analysis suggested that qigong has some particular characteristics which differentiate it from close-eyed rest and sleep state. In the tranquility state which is induced by mind and breath regulation, the specific activation of the cerebral cortex, the low frequency oscillation rhythm and harmonic hypothalamic-pituitary-andrenocortical regulation are likely to enhance the immunogenicity, facilitate the cerebral blood flow and microcirculation of the body. High levels of concentration, the "let go" attitude and positive thoughts of the "mind regulation" further enhance positive personality traits and delays intelligence decline. Qigong, therefore, seems to satisfy needs related to autonomy, competence and cognitive ability. This analytical review provided more positive evidence supporting the benefits of qigong on health, more specifically on reviewing the neurophysiological mechanism behind. Qigong can be a viable alternative to other intensive physical activities to achieve body-mind wellbeing.

Unfortunately, two electrodermal studies were excluded for further analysis as no control group was used in these two studies. Though findings indicated an increase in radiation and luminous discharges from the participants' hands, the data was inadequate for us to analyze radiation and the luminous phenomenon of qigong in an in-depth manner which needs further effort (Rubik and Brooks, 2005; Fan, 1990). Due to the special nature of EEG and fMRI measurements, it is common to have small sample size in a study. Most of the time, double blinding and randomization were not quite possible. This analysis had some methodological limitations and included some studies of less sophisticated design (e.g. non-RCT). With limited availability of RCT

neurophysiological studies and a variety of qigong forms, the conclusions should be treated with some caution. Different forms of qigong might have different effects on the participants. Further investigations on different styles of qigong might help to provide more evidence of the benefits of qigong in a more comprehensive manner.

Under the influence of western culture, qigong is sometimes viewed as hypnosis (Lee, 2000; Vickers and Zollman ,1999). However, early in the 1950's, it was found that although theta waves were present in both hypnotic and qigong states, the former occurred under the absence of alpha while the later happened when alpha waves were still present (HeBCKEY, 1958). Theta and alpha oscillations are regarded as multifunctional neuronal networks, reflecting attention, affective and cognitive activity inside the brain (Aftanas and Golocheikin, 2001). Tao, Zhang and Chen (2005) further illustrated that sensory evoked potentials (VEP, AEP) presented low amplitude in hypnotic states which indicated an obvious inhibition of visual and auditory sensory areas. Similar reactions were not present during qigong meditation. Zhang, Zhao, et al. (1988) showed that the frequency of EEG in TM was rather regular and demonstrated larger amplitude. However, alpha rhythm was not that stable and demonstrated some unusual excitations throughout the qigong practice. The findings suggested that the neurological impact of qigong was different from other meditative states. There might be a complex attention and emotional processing. However, these few studies involving both qigong and other mindful exercises (i.e. hypnosis or TM) employed within group design only, thus without using control group for comparison. Therefore, they were excluded in the first screening process for further analysis. Future sophisticated research effort may help to distinguish qigong from other mindful exercises in a more scientific and neurological aspect.

We suggest future work to test the qigong mechanism in a longitudinal, RCT and double-blind setting using a standardized protocol and validated measurements. Humans are very complex beings functioning on many different levels. It is suggested that wellbeing should be viewed from a biopsychosocial aspect (Chow and Tsang, 2007; WHO, 2001). However, most available studies used either psychological measuring tools (e.g. inventory) or physiological instruments (e.g. EEG) but not both. It is hard to gain a comprehensive picture of body-mind wellness. Therefore, we suggest that both psychological and physiological measurements should be included in future studies so as to investigate the holistic effects of qigong on body-mind wellbeing.

REFERENCES

Acuna C, Bolis M. Stigmatization and access to health care in Latin America: Challenges and perspectives. Pan American Health Organization, World Health Organization, 2005. Retrieved on May 25, 2011, from: http://www.paho.org/English/DPM/SHD/HP/leg-stigmatization-eng.pdf.

Aftanas LI, Golocheikine SA. Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalize attention: high-resolution EEG investigation of mediation. *Neurosci. Lett.* 2001; 310:57-60.

Baum A, Posuszny DM. Health psychology : mapping bio-behavioural contributions to health and illness. *Annu. Rev. Psychol.* 1999; 50:137-163.

Beehr TA. Perceived situation moderators of the relationship between subjective role ambiguity and role strain. *J. Appl. Psychol.* 1976; 61: 35-40.

Beehr TA, Walsh JT, Taber TD. Relationships of stress to individually and organizationally valued states: Higher order needs as a moderator. *J. Appl. Psychol.* 1976; 61: 35-40.

Berger BG, Owen DR. Mood alternation with yoga and swimming: Aerobic exercise may not be necessary. *Percept Mot. Skills* 1992; 75: 1331–1343.

Berger BG, Owen DR. Stress reduction and mood enhancement in four exercise modes: Swimming, body conditioning, Hatha yoga, and fencing. *Res. Q. Exerc. Sport* 1988; 59: 148–159.

Bernardi L, Porta C, Spicuzza L, Bellwon J, Spadacini G, Frey AW, et al. Slow breathing increases arterial baroreflex sensitivity in patients with chronic heart failure. *Circulation* 2002; 105 (2):143–145.

Bloomfield MD. Healing Anxiety with Herbs. New York: Harper Collins, 1998.

Brekke M, Hjortdahl P, Kvien TK. Changes in self-efficacy and health status over 5 years: A longitudinal observational study of 306 patients with rheumatoid arthritis. *Arthritis Care Res.* 2003; 49: 342-348.

Brosschot J F, Gerin W, Thayer J F. The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *J. Psychosom. Res.* 2006; 60: 113–124.

Brown RP, Gerbarg PL. Sudarshan Kriya yoga breathing in the treatment of stress, anxiety and depression: Part I – Neurophysiologic model. *J. Altern. Complement Med.* 2005; 11 (1): 189-201.

Byrne DG, Espnes GA. Occupational stress and cardiovascular disease. *Stress Health* 2008; 24: 231-238.

Cahn BR, Polich J. Meditation states and traits: EEG, ERP and neuroimaging studies. *Psychol. Bull* 2006; 132 (2): 180-211.

Caplan RD, Cobb S, French, JRP, Jr. Van Harrison R, Pinneau SR. Job Demands and Worker Health: Main Effects and Occupational Differences. Washington DC: US Government Printing Office, 1975.

Census and Statistics Department, Hong Kong. Persons with disabilities and chronic diseases. Special Topics Report No.48, Hong Kong: Census and Statistics Department, 2008.

Chan AS, Cheung MC, Sze SL, Leung WWM, D Shi. Shaolin dan tian breathing fosters relaxed and attentive mind: a randomized controlled neuro-electrophysiological study. *Evid. base Compl. Alternative Med.* 2011; doi:10.1155/2011/180704.

Chen K. An analytic review of studies on measuring effects of external qi in China. *Alternative Ther. Health Med.* 2004; 10: 38-50.

Chow YWY, Dorcas A, Siu A. The Effects of qigong on reducing stress and anxiety and enhancing body-mind well-being. *Mindfulness*, 2012; 3:51-59.

Chow YWY, Tsang HWH. Biopsychosocial effects of qigong as a mindful exercise for people with anxiety disorders: A speculative review. *J. Altern Complement Med.* 2007; 13(8): 831-840

Cobrun D. Job-worker incongruence: Consequences for health. *J. Health Soc. Behav.* 1975; 16,198-212.

Collins M, Dunn L. The effects of meditation and visual imagery on an immune system disorder: dermatomyositis. *J. Altern. Complement Med* .2005; 11: 275-84.

Croog SH, Levine S, Testa MA, Marcia A, Brown B, et al. The effects of antihypertensive therapy on the quality of life. *N. Engl. J. Med.* 1986; 314: 1657-1664.

Daley AJ. Exercise therapy and mental health in clinical populations: Is exercise therapy a worthwhile intervention? *Adv. Psychiatr. Treat.* 2002; 8: 262-270.

Davidson RJ, Kabat-Zinn J, Schumacher J, Rosenkranz M, Muller D, Santorelli SF, et al. Alterations in brain and immune function produced by mindfulness meditation. *Psychosom Med.* 2003; 65: 564–70.

De Jong MJ, Moser DK, An K, Chung M L. Anxiety is not manifested by elevated heart rate and blood pressure in acutely ill cardiac patients. *Eur. J. Cardiovasc. Nurs.* 2004; 3: 247-253.

Denollet J, Brutsaert DL. Personality, disease severity, and the risk of long-term cardiac events in patients with a decreased ejection fraction after myocardial infarction. *Circulation* 1998; 97:167–73.

Durham NC. Moderate exercise yields cardiovascular benefits. Science Daily, October 10, 2005. Retrieved on 09 December 2006, from: http://www.dukehealth.org/news/9296.

Fan JJ, Li SB, Wang F, Yang ZK, Yan HC, Li JS, Hou SJ, et al. Investigations of electromagnetic radiation and luminous phenomena in the state of qigong function [in Chinese]. *J. Fourth Milit Med.* Univ (*Di 4 jun yi da xue xue bao*) 1990; 11(2):149-150.

Fang L, Schinke SP. Complementary alternative medicine use among Chinese Americans: Findings from a community mental health service population. *Psychiatric Services* 2007; 58 (3): 402-404.

Fong DY, Hung TM, Huang YC. The effects of qigong on relaxation and emotion. *Bulletin of Sports and Exercise Psychology of Taiwan* 2004; 5: 19-42.

Forge R L. Mind-body fitness: Encouraging prospects for primary and secondary prevention. *J. Cardiovasc. Nurs.* 1997; 11: 53-65.

Forge RL. Aligning mind and body: Exploring the disciplines of mindful exercise. *ACMS's Health and Fitness Journal* 2005; 9(5): 7-14.

Gemmil GR, Heisler WJ. Machiavellianism as a factor in managerial job strain, job satisfaction, and upward mobility. *Acad. Manage J.* 1972*;* 15:51-62.

Goldberg D, Huxley P. Common Mental Disorders. London: Routledge, 1992.

Grevier M. CBT can be therapeutic: using cognitive behavioural therapy in primary care. *Informed*, 2003; 9: 4-5.

Guo XF, Zhou B, Nishimura T, Teramukai S, Fukushima M. Clinical effect of qigong practice on essential hypertension: a meta-analysis of randomized control trials. *The J. Altern Complement Med.* 2008; 14(1):27-37.

Hashimoto M, Mukasa H, Yamada S, Nakamura J, Inanaga K. Frontal midline theta activity and platelet MAO in human subjects. *Biol. Psychiatry* 1988; 23:31-43.

Haslam C, Brown S, Atkinson S, Haslam R. Patients' experiences of medication for anxiety and depression: Effects on working life. *Fam. Pract.* 2003; 21: 204-212.

He HZ, Li DL, Xi WB, Zhang CL. A "stress meter" assessment of the degree of relaxation in qigong versus non-qigong meditation. *Frontier Perspectives*, 1999; 8:37-42.

HeBCKEY MP (НеВСКИЙ МП). In Psychotherapy (Психотерапии) [in Cyrillic], Medgee (Медгиз) , 1958, 63.

Henderson RJ, Hart MG, Lai SK, Hunyor SN. The effect of home training with direct blood pressure biofeedback of hypertensives: A placebo-controlled study. *J. Hypertens* 1998; 16: 771-778.

Inanaga K. Frontal midline theta rhythm and mental activity. *Psychiatry Clin. Neurosci* 1998; 52: 555-566.

Kawano K, Koita H, Fujiki T, Shinagawa Y. EEG and topography during Chinese "qigong" training. *Neurosciences* 1990; 16 : 503-508.

Kessler D, Lloyd K, Lewis G. Cross sectional study of symptom attribution and recognition of depression and anxiety in primary care. *Br. Med. J.* 1999; 318: 436–440.

Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Arch. Gen. Psychiatr.* 2005; 62: 617–27.

Kokubo H. Concept of 'Qi' or 'Ki' in Japanese qigong research. Proceedings of Present Papers of the 44th Annual Convention of the Parapsyhcological Association 2001; 147-154.

Konstam V, Moser D K, De Jong M. Depression and anxiety in heart failure. *J. Card Fail* 2005; 11: 455-463.

Kuo TC, Ho FM, Lin CW. Physiological changes associated with Neiyang qigong. *Journal of Health Science*, 2003; 49: 278-284.

Lan C, Chou SW, Chen SY, Lai JS, Wong MK. The aerobic capacity and ventilatory efficiency during exercise in qigong and tai chi chuan practitioners. *Am. J. Chin. Med.* 2004; 32 (1): 141-150.

Lan C, Chou SW, Chen SY, Lai JS, Wong MK. The aerobic capacity and ventilatory efficiency during exercise in qigong and tai chi chuan practitioners. *Am. J. Chin. Med.* 2004; 32: 141-150.

Lee MS, Huh H J, Jeong SM, Lee HS, Ryu H, Park JH, et al. Effects of Qigong on immune cells. *Am. J. Chin. Med.* 2003; 31: 327–35.

Lee MS, Huh HJ, Hong SS, Jang HS, Ryu H, Lee HS, et al. Psychoneuroimmunological effects of Qi-therapy : preliminary study on the changes of level of anxiety, mood, cortisol and melatonin and cellular function of neutrophil and natural killer cells. *Stress Health* 2001; 17: 17-24.

Lee MS, Huh HJ, Kim BG, Ryu H, Lee HS, Kim JM, et al. Effects of qi-training on heart rate variability. *Am. J. Chin. Med.* 2002; 30: 463-470.

Lee MS, Kang CW, Lim HJ, Lee MS. Effects of qi-training on anxiety and plasma concentrations of cortisol, ACTH, and aldosterone: A randomized placebo-controlled pilot study. *Stress Health* 2004; 20: 243-248.

Lee MS, Kim BG, Huh HJ, Ryu H, Lee HS, Chung HT. Effects of qi-training on blood pressure, heart rate and respiration rate. *Clin. Physiol.* 2000; 20 (3): 173-176.

Lee MS, Lee MS, Kim HJ, Moon SR. Qigong reduced blood pressure and catecholamine levels of patients with essential hypertension. *Int. J. Neurosci.* 2003; 113: 1691-1701.

Lee MS, Lim HJ, Lee MS. Impact of qigong exercise on self-efficacy and other cognitive perceptual variables in patients with essential hypertension. *J. Altern. Complement. Med.* 2004;10: 675-680.

Lee MS, Ryu H, Chung HT. Stress management by psychosomatic training: Effects of ChunDoSunBup qi-training on symptoms of stress: A cross sectional study. *Stress Med.* 2000; 16 (3): 161-166.

Lee MS. Qi therapy as an intervention to reduce chronic pain and to enhance mood in elderly subjects: A pilot study. *Am. J. Chin. Med.* 2001; 29: 237-245.

Lee S, Lee M, Chiu M, Kleinman A. Experience of social stigma by people with schizophrenia in Hong Kong. *Br. J. Psychiatr.* 2005; 186: 153–157.

Lee S. Chinese hypnosis can cause qigong induced mental disorders. *Br. Med. J.* 2000; 320: 803.

Lee SH, Choi TK, Yook KW. Effectiveness of a qigong-based stress management program in patients with anxiety disorder. *Journal of the Korean Neuropsychiatric Association* 2005; 44 (6): 690-699.

Lee SL, Bae BH, Ryu H, Sohn JH, Kim SY, Chung HT. Changes in alpha wave and state anxiety during ChunDoSunBup qi-training in trainees with open eyes. *Am. J. Chin. Med.* 1997; 27 (3-4): 289-299.

Lehrer P M, Woolfolk R L, Sime WE, eds. Principles and Practice of Stress Management, 3rd ed. New York: The Guilford Press, 2007.

Leung TH, Au W, Ching R, Heung L, Lam C, Leung E, et al. (eds.). Non-Communicable Disease Watch, 2009, 2. Retrieved on Sept 8, 2011, from: http://www.chp.gov.hk/files/pdf/NCD_Watch_Aug_09_Eng.pdf.

Leung Y, Singhal A. An examination of the relationship between qigong meditation and personality. *Soc. Behav. Pers.* 2004; 32: 313-320.

Li BP. Tu jie Zhongguo Yang Sheng Shu (圖解中國養生術) [in Chinese]. Taiwan: Hua wei guo ji, 2009.

Lim YA, Boone T, Flarity JR, Thompson WR. Effects of qigong on cardiorespiratory changes: A preliminary study. *Am. J. Chin. Med.* 1993; 21 (1): 1-6.

Lindforts P, Lundberg U. Is low cortisol release an indicator of positive health? *Stress Health*, 2002; 18: 152–160.

Liu BH, Jiao J, Li YF. The effect of qigong exercise on the blood level of monoamine neuro-transmitters in the patients with chronic diseases [in Chinese]. *Chin J. Integr. Med.* 1990; 10(4): 195,203-206.

Lovallo WR. Stress and Health: Biological and Psychological Interactions, 2nd ed. Thousand Oaks, California: SAGE, 2005.

Mackenzie ER, Rakel B, eds. Complementary and Alternative Medicine for Older Adults. New York: Springer Publishing Company, 2006.

McEwen BS. Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators. *Eur. J. Pharmacol.*, 2008; 583(2-3):174-85.

McKinney CH, Antoni MH, Kumar M, Tims FC, McCabe PM. Effects of guided imagery and music (GIM) therapy on mood and cortisol in healthy adults. *Health Psychol.* 1997; 16, 390–400.

Netz Y, Lidor R. Mood alterations in mindful versus aerobic exercise modes. *J. Psychol.*, 2003; 137: 405–419.

Ng BHP, Tsang HWH. Psychophysiological outcomes of health qigong for chronic conditions: A systematic review. *Psychophysiology* 2009; 46 (2):257-269.

Ng D M, Jeffery RW. Relationships between perceived stress and health behaviours in a sample of working adults. *Health Psychol.* 2003; 22: 638–642.

Pan WX, Zhang LF, Xian Y. The difference in EEG theta waves between concentrative and non-concentrative qigong states –a power spectrum and topographic mapping study. *J. Tradit. Chin. Med.* 1994; 14(3): 212-218.

Pandya D. Psychological stress, emotional behaviour and coronary heart disease. *Compr. Ther.* 1998; 24: 265-271.

Pavek RR. Effects of qigong on psychosomatic and other emotionally rooted disorders. Proceedings of First World Conference of Academic Exchange on Medical Qigong, Beijing, China, 1988; 150.

Qin Z, Jin Y, Lin S, Hermanowicz NS. A forty-five follow-up EEG study of qigong practice. *Int. J. Neurosci.* 2009; 119: 538-552.

Qin Z, Zhai ZP. EEG studies of qigong [in Chinese]. *Shanghai Journal of Traditional Chinese Medicine* (*Shanghai zhongyiyao zazhi*) 1962; 5: 10-13, 27.

Rose S, Taylor P. Selected characteristics of people on long-term sick leave. *Psychiatr Bull* 1998; 22: 432–435.

Rubik B, Brooks AJ. Digital high-voltage electrophotographic measures of fingertips of subjects pre-and post-qigong. *Evid. Base Integr. Med.* 2005; 2(4):245-252.

Schell FJ, Allolio B, Schonecke OW. Physiological and psychological effects of Hatha-yoga exercise in healthy women. *Int. J. Psychosom.* 1994; 41: 46–52.

Schneider RH, Staggers F, Alexander CN, Sheppard W, Rainforth M, Kondwani K., et al. A randomised controlled trial of stress reduction for hypertension in older African Americans. *Hypertension* 1995; 26: 820–827.

Scully D, Kremer J, Meade MM, Graham R, Dudgeon K. Physical exercise and psychological well being: A critical review. *Br. J. Sports Med.* 1998; 32: 111–120.

Segail R. A Japanese killer: Stress in Japan has become a national killer. Psychology Today Magazine, 2007; Sept/Oct. Retrieved on December 15, 2011, from: http://www.psychologytoday.com/articles/200009/japanses-killer.

Shan H, Yan H, Sheng H, Hu S. A preliminary evaluation on Chinese qigong treatment of anxiety. *Proceedings of the Second International Conference on Qigong*, Xian, China, 1989; 165.

Shen ZY, Shinlin, Zhu YQ, Yu P, Zhang QH, Qu MH. Nonlinear analysis and research on multi-channel EEG phase-space reconstruction at qigong mediation [in Chinese]. *Acta Universitatis Traditionis Medicalis Sinensis Pharmacologiaeque Shanghai* 2008; 22 (6): 54-56.

Silva LMT, Cignolini A, Warren R, Budden S, Gooch AS. Improvement in sensory impairment and social interaction in young children with autism following treatment with an original qigong massage methodology. *Am. J. Chin. Med.* 2007; 35 (3): 393-406.

Skoglund L, Jansson E. Qigong reduces stress in computer operators. *Complement Ther. Clin. Pract.* 2007; 13: 78-84.

Stahl SM, Hauger R. Stress: an overview of the literature with emphasis on job-related strain and intervention. *Adv. Ther.* 1994; 11: 110-119.

Sun FL, Lei SP, Yan YA, He QN. Research on using qigong to delay reduction of intelligence in aged people [in Chinese]. *Acta Medica Sinica* (*Xin Li Xue Bao*) 1989; 4(3): 10-12,79.

Sun FL, Wang JS, Liu GZ, Jiao XY, Zhang ZJ, Shi YJ, Zhang TG. An analysis on EEG power spectrum and coherence during quiet state in qigong [in Chinese]. *Acta Psychologica Sinica* (*Xin Li Xue Bao*) 1984; 4:422-427.

Sundar S, Agrawal SK, Singh VP, Bhattacharya SK, Udupa KN, Vaish SK. Role of yoga in management of essential hypertension. *Acta Cardiol.* 1984; 39: 203–208.

Takahashi M, Brown S. Qigong for Health. New York: Japan Publication, 1988.

Takahashi T, Murata T, Hamada T, Omori M, Kosaka H, Kikuchi M, et al. Changes in EEG and autonomic nervous activity during meditation and their association with personality traits. *Int. J. Psychophysiol.* 2005; 55: 199–207.

Tang CM, Wei X, Zhang JP, Lu ZY, Wang JM, Liang F, Sun BY, et al. Effects of qigong on personality – A longitudinal study [in Chinese]. *Chinese Journal of Behavioral Medical Science* 1995; 4 (1): 15-18.

Tao MY, Zhang MD,Chen XS. Experiment of brain evoked potentials in "qigong state" and "hypnosis" of normal adults [in Chinese]. *Chin J. Clin. Rehabil* 2005; 9 (24): 255-269.

Tennant C. Work stress and coronary heart disease. *J. Cardiovasc. Risk* 2000; 7: 273-276.

Tsang HWH, Cheung L, Lak DCC. Qigong as a psychosocial intervention for depressed elderly with chronic physical illnesses. *Int. J. Geriatr. Psychiatr.* 2002; 17: 1146–1154.

Tsang HWH, Fung KMT, Chan ASM, Lee G, Chan F. Effect of a qigong exercise programme on elderly with depression. *Int. J. Geriatr. Psychiatr.* 2006; 21: 890 – 897.

Tsang HWH, Mok CK, Au Yeung YTA, Chan SYC. The effect of Qigong on general and psychosocial health of elderly with chronic physical illnesses: A randomized clinical trial. *Int. J. Geriatr. Psychiatr.* 2003; 18: 441–449.

Tseng WS, Lin TY , Yeh EK, eds. Chinese Societies and Mental Health. Hong Kong: Oxford University Press, 1995.

U.S. Department of Health and Human Services 2001. Health Report 2000. U.S.A. Retrieved on September 10, 2011 from: http://www.cdc.gov/nchs/data/hp2000/hp2k01.pdf.

Vickers A , Zollman C. ABC of complementary medicine – Hypnosis and relaxation therapies. *Br. Med. J.* 1999; 319:1346-1349.

Wang J. Role of qigong on mental health. Proceedings of the Second World Conference of Academic Exchange on Medical Qigong, Beijing, China. 1993.

West J, Otte C, Geher K, Johnson J, Mohr DC. Effects of Hatha yoga and African dance on perceived stress, affect and salivary cortisol. *Ann. Behav. Med.* 2004; 28: 114–118.

Wirht D, Cram JR, Chang RJ. Multisite electromyographic analysis of therapeutic touch and qigong therapy. *Altern Complement Med.* 1997; 3(2),109-118.

World Health Organization [WHO]. Basic Documents, 43rd ed. Geneva: WHO Press, 2001; 1.

Yao S. Qigong Yangsheng Fa. (Health Preservation Method of Qigong). Hong Kong: Bo Yi Publishing Co., 1991.

Yu P, Shen ZY, Chai JY. Heart rate variation in heart regulation of qigong. Acta Universitatis Traditionis Medicalis Sinensis Pharmacologiaeque Shanghai 2000; 14 (3): 40-41.

Yu WL, Li XQ, Tang WJ, Li Y, Li K, Weng XC, Chen YZ. Resting-state fMRI study of low frequency synchrony in motor cortex under the state of "qigong" [in Chinese]. *Chin Comput. Med. Imag.* 2005; 11(6):370-373.

Yuen H. The psychological effects of qigong practice [in Chinese]. Dong Fang Qigong 2000; 6: 2-6.

Zhang GF, Fu CF, Zhao YF. The dual-directional regulation of qigong on cardiovascular function [in Chinese]. *Heilongjiang Medicine and Pharmacy* 1990; 13(4): 324-325.

Zhang JZ, Li JZ, He QN. Statistical brain topographic mapping analysis for EEGs recorded during qigong state. *Intern J. Neuroscience* 1988; 38:415-425.

Zhang JZ, Zhao J, He QN. EEG findings during special psychical state (qigong state) by means of compressed spectral array and topographic mapping. *Comput. Biol. Med.* 1988; 18: 455-463.

Zhang TG. The origin of the name "qigong" ("氣功" 一詞的由來) [in Chinese]. *Xian Dai Yang Sheng* 2005; 8.

Zhang YH, Chen C, Xu JF. A technology based on multi-measure dynamic brain electric activity mapping and consciousness character in qigong practice [in Chinese]. *ACTA Universitatis Traditionis Medicalis Sinensis Pharmacologiaeque Shanghai* 2008; 22(4):54-57.

In: Philosophy of Mind ISBN: 978-1-62257-215-1
Editors: R. J. Jenkins, W. E. Sullivan

Chapter 4

THE CONSCIOUS MIND AND THE UNCONSCIOUS MIND: A DECISION THEORY ANALYSIS WITH IMPLICATIONS FOR SOCIETY, RELIGION AND DISEASE

James A. Morris*
Medical Undergraduate Education,
University Hospitals of Morecambe Bay
NHS Foundation Trust, Royal Lancaster Infirmary,
Lancaster, UK

ABSTRACT

Decisions are based on estimates of *a priori* probability, evidence and the values and costs of the anticipated outcome. An optimal decision strategy is one which seeks to maximize expected value. Digital computers, neural networks, robots, the unconscious brain and the conscious mind are all capable of following an optimal decision strategy. The rules by which values and costs are attached to specific outcomes are pre-programmed but then modified by experience even in automata. The conscious mind, however, must experience the values and costs and

* E-mail: Jim.A.Morris@mbht.nhs.uk

therefore it is directly rewarded or punished by the unconscious brain. This is the profound insight of psycho-analysis expressed in modern parlance. Molecules circulating in the blood act directly on the brain controlling mood and motivation. They switch on and off genetic and proteomic networks that control and modulate conscious experience. The rules of human behavior influenced by evolutionary genetics but extensively modified by social experience are written in the networks and used to reward or punish the conscious mind. These ideas have implications for normal social functioning, altruism, aesthetics, the determinants of happiness and the universal phenomenon of religion. But complex genetic and proteomic networks can malfunction and lead to disabling diseases such as schizophrenia, depression and functional psychiatric disorders. Faulty genetic networks caused by deleterious mutations might play a part, particularly in the etiology of schizophrenia, but the exciting possibility that microbial molecules play a major role is also explored. If a genetic or proteomic network is switched on or off without concomitant inflammation possibilities include bacterial toxins absorbed from the intestinal tract or antibodies directed against microbes which cross react with brain proteins.

We will improve our health once we appreciate that social interaction can lead to happiness through acts of altruism which generate conscious reward. Our spiritual life is enhanced by the realization that religion will inevitably emerge when intelligent conscious animals interact. There is a good chance we can reduce disease, including diseases of the mind, by controlling the rate, dose and route of microbial exposure so as to optimize the immune response and reduce the chance of infection and the generation of auto-antibodies. All this follows from a philosophy of the mind in which information theory and decision theory play a key role. This philosophy gives us a deeper understanding of the nature of consciousness and points to practical solutions to problems that currently trouble human kind.

Keywords: consciousness, decision theory, altruism, happiness, beauty, religion, depression, schizophrenia, functional psychiatric disorder, microbial toxins, molecular mimicry, auto-antibodies

INTRODUCTION

In the last one hundred years there have been enormous strides in understanding the nature of biology. The information that specifies an organism is coded in the genome and expressed in the proteome. In humankind there are approximately 25,000 genes but ten fold more proteins

[1]. The proteins determine structure and function; they act as carriers of information between cells and orchestrate the movement of ions across cell membranes.

Today we can claim to comprehend the nature of vitality; that which distinguishes the living from the dead and the animate from the inanimate. To describe vitality, however, requires much more than a single sentence, or a paragraph or even a chapter. It depends on an enormous amount of detail that has been acquired over many years. We can identify key landmarks on the road to understanding but there was no single moment when it all became clear.

I suspect the same will apply to the last great question in biology: what is the nature of consciousness? We will not answer the question directly as a result of a sudden single brilliant insight. Instead it will be the slow, steady acquisition of knowledge that characterizes scientific endeavor. Consciousness, like vitality, is an emergent property of a complex system and we will not be able to grasp its nature until we understand the full complexity of the biological systems that are conscious and those that are not.

A potentially valuable approach to the study of consciousness is to use insights from information theory and decision theory [2, 3]. This is the approach used in this chapter. Biological automata process information and make decisions and this can be analyzed in term of decision theory. An optimum decision strategy is one that maximizes expected value [4]. *A priori* probability, evidence and the benefits and costs of right and wrong decisions are computed using a common scale of measurement. Maximizing expected value is concerned with survival and reproduction. Conscious animals also make decisions but the values and costs are subjective sensations of pleasure and pain. The goal is to maximize expected pleasure and minimize expected pain. The rewards and punishments of pleasure and pain are delivered to the conscious mind by non conscious parts of our brain and body. The rules that govern this reward system are in turn determined in part by our genes and in part by what we learn during development. The moral rules that govern society are programmed in our subconscious and govern our conscious actions. The implications of this for morality, religion, individual happiness and psychological disease are explored.

In many ways the interaction of the conscious and the subconscious discussed in this chapter is a modern formulation of the profound insights of Freud, Jung and Adler - the giants of psychoanalytical theory. Their ideas, however, need to be integrated with information theory and with modern molecular biology. The genome and the proteome underpin conscious and

subconscious thought and we will not have a complete understanding until we know the molecular biology in detail.

DECISIONS

Biological organisms react to a potentially threatening world so as to preserve their integrity for long enough to pass on their genes to the next generation. This process can be analyzed in terms of decision theory. There is a flow of information from the outside, a process of analysis and then a decision between alternative courses of action designed to meet the organism's goals. This applies to all animals; and in humankind it applies to our conscious decisions as well as decisions made at a non-conscious level.

The aim in classical decision theory is to maximize expected value [4] (table). The process involves taking into account *a priori* probability, evidence gathered from the environment and the benefits and costs associated with different courses of action. Let us consider a specific example.

I have practiced as a consultant pathologist in the NHS for over 40 years. The job of a pathologist is laboratory based diagnosis of disease. I have a blood test for disease A which is 95% sensitive and 99% specific. This means that if the patient suffers from disease A my test will be positive 95% of the time and negative 5% of the time.

If the patient does not suffer from disease A then the test will be negative 99% of the time and positive 1% of the time. Unfortunately in pathology, as in the rest of life, we never achieve certainty. The consultant physician who sent the blood specimen to the laboratory for analysis is very good at diagnosing disease A. She always takes a detailed history and does a careful examination of the patient. In this specific case she says there is a 50% chance that the patient does have the disease. As far as I am concerned this is the starting probability i.e the *a priori* probability. It is 50% or 0.5 expressed as a probability.

If the blood test is positive the probability that the patient has the disease is 0.989 and if the blood test is negative the probability that the patient does not have the disease is 0.951. Those who like mathematics can work this out for themselves. Let us imagine there are 200 patients; 100 have the disease and 100 do not (*a priori* probability 0.5). If I test all 200 there will be 95 patients who have the disease and test positive, 5 who have the disease and test negative, 99 who do not have the disease and test negative and 1 who does not have the disease and tests positive.

Table. A decision which maximizes expected value depends on *a priori* probability, evidence and the values and costs associated with right and wrong decisions. Values and costs need to be expressed in the same units on a one dimensional scale, and the validity of measures of happiness indicates that this can be done in practice

Decision theory Consider two possible hypotheses h_1 and h_0 With correct responses H_1 given h_1 and H_0 given h_0 The response that maximizes expected value is to respond H_1 when: $L_{10}(e) > B$ Where $L_{10} = P[e/h_1]$ divided by $P[e/h_0]$ And $B = (V_{00} + V_{01})(P[h_0])$ divided by $(V_{11} + V_{10})(P[h_1])$ L_{10} is the likelihood ratio of event *e* for h_1 relative to h_0 $P[e/h_1]$ is the probability of event *e* given h_1 $P[e/h_0]$ is the probability of event *e* given h_0 $P[h_1]$ is the *a priori* probability of h_1 $P[h_0]$ is the *a priori* probability of h_0 V_{00} is the value of response H_0 given h_0 V_{01} is the cost of the response H_1 given h_0 V_{11} is the value of the response H_1given h_1 V_{10} is the cost of the response H_0 given h_1

The positive predicted probability is 95 divided by 96 (= 0.989) i.e of those who test positive 95 have the disease and 1 does not. The negative predictive probability is 99 divided by 104 (= 0.951) i.e of those who test negative 99 do not have the disease but 5 do.

Let us now repeat the calculation when the *a priori* probability is 0.05. In this situation the physician states that the patient is worried that he has disease A but she thinks it unlikely and would estimate the chance as 5% rather than 50%. The positive predictive probability is now 0.83 and the negative predictive probability is 0.997. If the test is negative the patient can be reassured but if the test is positive there is a problem. Thus the interpretation of the test depends not only on the result but also on *a priori* probability.

However in medicine the next stage is to consider treatment. The treatment for disease A is drug *A*. If disease A is potentially fatal and drug *A* has few side effects we will recommend treatment to the patient whose positive predictive probability is only 0.83. But if disease A is less severe and drug *A* has a risk of serious side effects we will be less inclined to treat the patient with a positive predictive value of 0.83 but we might still treat the patient with a positive predictive probability of 0.989. To make a precise calculation we need to determine the benefit of treatment to the patient who does have the disease and the cost of treatment to the patient who does not have the disease and then set that against the probability that the patient does or does not have the disease. Furthermore the benefits and costs need to be expressed in the same units (table).

In practice we rarely have precise estimates of *a priori* probability. We often do not know the precise sensitivity and specificity of pathological tests and it is extremely difficult to assess the benefits and costs of treatment and express them in the same units. But it is clear that we need to take each of the three elements into consideration or our medical decisions will be far from optimal. It is also clear that the same general principles must apply to other decisions that we make at a conscious or non conscious level. Decisions must consider *a priori* probability, evidence and the benefits and costs associated with the actions that arise from the decision.

Let us now consider a different scenario. Tom is placed in the middle of a large room with two doors; one at either end [5]. One door is Tom's potential exit. The other leads to another room in which there is an animal. Tom is told that there is a 0.5 probability that the animal is a man eating tiger, and a 0.5 probability that it is a man holding a cheque for $1,000,000 made out in Tom's name. He is also told that in 5 minutes time the door will be opened and the animal will emerge. If the tiger emerges and Tom is still present he will be killed. If Tom leaves before the man with the cheque emerges he will not get his money. What should Tom do? Well first of all he should spend 4 minutes 30 seconds trying to work out if there is any way that he can determine the nature of the animal before the door opens. If not then a rational decision

depends on the value he places on his life compared with the value of the cheque. If Tom has a nice life and an adequate income he will leave. But if he is a tramp living an unpleasant life sleeping on the streets he might be prepared to risk his life for $1,000,000. But if Tom was told that the probability that the animal was a man eating tiger was only 0.001 and the probability that it was a man with a cheque for $1,000,000 was 0.999 the calculation is different. Many, perhaps most, would take the risk. There is only a 1 in a thousand chance of losing one's life against a near certainty of a million dollar payment. The mental process involves expressing the value of one's life and the value of the cheque in the same units and then working out the expected gain or loss depending on the *a priori* probability. Individuals do not actually perform the calculations in their mind but the process that leads to the decision must be in some sense equivalent.

A further modification would be to place a partially opaque window between the rooms in which Tom was placed and the room in which the animal was placed. Tom would then have 4 minutes 30 seconds in which to try and determine the nature of the animal. The amount of evidence he manages to gather by looking through the window will now influence his decision. If he is certain it is a tiger he will go, if he thinks it is a man he might stay. Once again the rational process is to integrate *a priori* probability, evidence and the benefits and costs of the alternative options of staying or going.

A long standing debate in psychology concerns the extent to which human beings follow a decision strategy that maximizes expected value [6]. A classical example is voting. I remember going out in the pouring rain to vote in a national general election. The probability that my vote could make a difference was minimal. The down side of turning out in the pouring rain was considerable. How can that action be rationale? The costs outweigh the expected benefits. But that ignores my moral qualms. I believe that voting is important. If I didn't vote I would feel bad. If I do vote I feel good and if I go out in the pouring rain to vote I feel especially good. The benefits outweigh the costs because the benefits to my conscious self are the emotional feelings that I experience. Those feelings are delivered to my conscious self by the rest of my brain and my body i.e by my non conscious self.

Another classical example is to offer a colleague a bet on the toss of a coin; $100 loss on heads against $200 gain on tails [6]. It is claimed that the rational decision is to accept the bet because the expected gain exceeds the expected loss. Most people do not accept the bet and therefore it is claimed that most people are not rationale in decision making; they do not follow an optimum decision strategy. But in fact the $100 loss is a much bigger loss to

most people than the $200 dollar gain. If the $100 in your wallet is what you need for food and transport over the next few days then the loss would be considerable. By comparison the $200 gain would have little or no effect on your life over the next few days or weeks. In addition if you regard a $100 loss as considerable the loss to your colleague of $200 would be even greater. You would feel bad if you lost the toss and even worse if you won. The rationale response, the one that maximizes expected value, is to refuse the bet.

NON CONSCIOUS DECISIONS

The task of developing an immune response to micro-organisms requires a vast number of decisions in which protein molecules are classified as self or not self [7]. The individual then creates clones of lymphocytes which are capable of recognizing and responding to foreign proteins on bacteria and viruses. These effector systems are capable of preventing infection by attacking foreign organisms but they avoid attacking our own cells.

Human beings have approximately 25,000 genes which code for approximately 250,000 proteins [1]. Each human gene has several exons and the RNA transcripts of the exons can be combined in various ways to produce several different proteins. In some cases exons from different genes can be combined at random to produce a large number of protein variants and this underlies the generation of antibody diversity and the diversity of the T lymphocyte receptors. In addition proteins can catalyze the transfer of chemical groups producing further diversity. Hence there are many more proteins in the proteome than there are genes in the genome.

Bacteria have 3000 to 5000 genes coding for 3000 to 5000 proteins. But there are of the order of 400 to 800 different types of bacteria that can be found on our mucosal surfaces with perhaps 40 to 80 present at any one time [8]. Viruses are simpler with fewer proteins but we meet hundreds if not thousands in a lifetime. Thus the total complexity of the microbial proteome to which we are exposed is similar in magnitude to our own proteome. Furthermore genes and their protein transcripts are conserved in evolution and it is no exaggeration to say that we share our proteome with the rest of creation.

In the initial phase of an immune response microbes are broken down within macrophages and their proteins are cut into polypeptides approximately 8 to 12 amino acids long.

There are 20 different amino acids, so the number of polypeptides 10 amino acids long is 20^{10}, which is approximately equal to 10^{13}.

There are far fewer than ten trillion proteins so each decapeptide is specific for the protein from which it was cut. The polypeptides of length 8 to 12 amino acids are combined with a molecule in the macrophage cytoplasm called MHC2 and the complex is then posted on the macrophage surface. T lymphocytes have a wide range of receptors and those that recognize and combine with the MHC2/polypeptide complex undergo clonal proliferation. The clone of cells is then specific for the particular polypeptide and therefore specific for the protein from which it was cut.

The T lymphocytes then recruit B lymphocytes which manufacture antibody which specifically recognizes parts of the protein molecule that the T cell has identified. Some of the T cell clones will recognize foreign proteins and some will recognize self proteins i.e proteins that have similar shapes to human proteins.

The immune system must in some way assess the potential for harm of each bacterium and virus it meets. It must then produce antibody to epitopes on foreign proteins that do not cross react with exposed self proteins. The clones of lymphocytes need to be concentrated on potentially pathogenic microbes rather than harmless commensals.

All of this occurs automatically, it is not under conscious control, but it nevertheless demands a high level of information processing and decision making. We do not understand every detail but we know enough to appreciate its complexity.

There are a number of general principles that apply when information is processed [2, 5, 9]:

1. Information is processed in noise and there is a finite chance of error.
2. Information processing systems have a finite capacity
3. Information processing systems decay at random with time according to the laws of entropy and the risk of error will rise with age.
4. Complex information processing systems require a high level of redundancy to reduce error to a minimum.

These rules can be used to gain insight into the process of generating an immune response.

Rule one: Information is processed in noise and there is a finite chance of error.

Let us define two possible types of error:

Type one error: a failure to classify a foreign protein as foreign and therefore a failure to produce an effective T and B cell clonal response to the protein. This will increase the risk of infection with that organism. This is a false negative.

Type two error: a failure to recognize that a foreign protein cross reacts with an exposed self protein. This means the development of a T and B cell response to exposed self proteins leading to autoimmune disease. This is a false positive.

Rule two: Information processing systems have a finite capacity.

If we are exposed to a large dose of a potentially pathogenic organism then a rapid response is required and this will increase the risk of both type one and type two errors. The best way to meet microbes is in low dose on the mucosal surface where the lymphocytic analytical machinery is concentrated.

Rule three: Information processing systems decay at random with time according to the laws of entropy and the risk of error will rise with age.

It's best to meet microbes early in life rather than later in life [9].

An interesting aspect of decision theory is that changing the decision criteria so as to reduce false positives will automatically increase false negatives and *vice versa.* Thus we should be able to recognize groups in the community who have increased risk of infection (type one error) but a decreased risk of autoimmune disease (type 2 error), with the complementary group having increased risk of autoimmune disease and decreased risk of infection. The mutually complementary and non-overlapping groups which show this are males and females [9, 10]. Men are more likely to die of infection than women at all ages. Women, particularly in middle life, are more likely to suffer from autoimmune disease than are men.

CONSCIOUS DECISIONS

Automata are capable of complex information processing and complex decision making. This applies to digital computers, neural networks and robots as well as the non conscious activities of biological organisms. Few of us will generate by conscious thought anything that can match the activities of our

own immune system. Thus whatever is the biological reason for the evolution of consciousness it is not complexity of decision making.

Let us consider a very clever robot. It can do the housework and the gardening and play a number of games, including chess, to a high level. Its cleverness depends on the sophistication of programming. There will be rules to guide the robot in the house and in the garden. It will have targets to reach in terms of tidiness in the house and standards for the garden. It will no doubt be very expensive to buy and so it will have strict rules to avoid any damage to itself. The robot will have neural networks so that it can learn from its environment but it will still be a difficult task to write programs that avoid all possible hazards.

Let us now take a leap of imagination and consider a robot which is conscious. There could be an advantage if it has some freedom to make decisions that are not purely automatic. It might be able to learn to deal with new hazards that were not anticipated in the initial programs. It might be able to extend its repertoire in the house and garden. It might come up with new games to play. I doubt that it will be more clever, but it could be more original and more adaptable. But how would it frame a conscious decision and what do benefits and costs mean to a conscious robot? Why would it want to keep the house tidy? Why would it want to avoid a hazard? What is to stop it throwing itself off a cliff? If the robot is conscious we would need to program the non conscious part to deliver rewards and punishments to the conscious part. The conscious part would then make decisions to maximize expected value by avoiding pain and seeking pleasure.

The same principle applies to biological organisms. The basic information that governs structure and function is stored in the genome and expressed in the proteome. The information was written many generations ago and there is no way it can anticipate every possible environment and every possible hazard. We learn from our environment using neural networks and we can modify our behavior to suit the circumstances. That implies some degree of freedom in decision making. But decisions can only be made if there are benefits and costs. To an automaton benefits and costs are programmed as numbers. But to a conscious animal benefits and costs are experienced as sensations and emotions. The conscious animal seeks to maximize pleasure and minimize pain. The rules that determine what is pleasurable and what is painful are determined by the non conscious part.

Most of the information that we process and most of the actions that follow from the decisions made occur at a subconscious or non conscious level. The information that reaches consciousness has already been extensively

processed. The conscious level is analogous to the board level of a company. It is presented with processed data giving an overview of activity. The board then decide overall strategy which should be consistent with the company goals. If the strategy is successful the company makes profits and the board members, as well as the employees, are rewarded. If the strategy is not successful they are both punished (at least that is the theory of capitalism). In the case of biological organisms the rewards and punishments are immediate and effective.

LANGUAGE

Language is species specific behavior. The information to enable us to understand and speak a language is coded in the genes but then extensively modified by experience. We are genetically programmed to speak but the language we learn depends on where we are born. Genes code for neural networks that interpret auditory information and control the organs of speech. In some sense there is a deep structure and rules of grammar that apply to all languages [11]. There is a great deal they have in common. But the words and some of the rules by which they are combined differ between different languages.

It is a common observation that children learn languages easily and quickly but this ability seems to fade with age. There are key periods in development when our neural networks are attuned to language acquisition. But in secondary school most of us struggle to learn a new language and in later life few can become fluent in a new language.

The evolutionary advantage to a social animal of language acquisition is obvious. Societies will be more coherent and function more effectively. Those who live in the society will enjoy better housing, better transport, a better food supply and be better protected against life's hazards. They are more likely to survive, marry, reproduce and have children. In the struggle to survive language is a definite plus for humankind.

It is possible to talk about the English language or the French language or any other language of a group in the abstract. The English language is an entity with a lexicon of words and rules of grammar even though no one person knows all the words or is consistent in applying the rules of grammar. We have different accents, a different range of words, and different styles of sentence construction. Furthermore language changes as new words enter the lexicon and old words acquire new meanings. Thus the English language is a

gradually changing abstract concept. It is not the language of any one person, not even of the Queen. It is the language of the group. It is an emergent property of the group.

MORALS

Human beings are social animals and society needs rules. Just as language is written in the genes and then modified by the environment so is morality. There is a common thread of morality which is found in all social groups analogous to the deep structure of language. But children learn the specific rules that apply to the society in which they are born and raised. I suspect there are key developmental periods in the acquisition of morality as there are in language but the evidence for this is less clear cut. Moral rules also gradually change with time and individuals differ in the extent to which they adhere to moral rules. But we recognize an overall moral code that belongs to the group even though it is not the precise moral code of any one person. It is an emergent property of the group.

The acquisition of rules, be they rules of grammar or morality, involve physical and chemical changes in neural networks in the brain. There is a structural basis to morality and although we are conscious of at least parts of our moral code most of the decisions are being made at a non conscious level. If the conscious mind decides on a course of action which offends against the rules of morality then the non-conscious mind will punish the conscious with feelings of guilt. This is not something the conscious mind can override. Equally if the conscious mind embarks on a course of action that fits with the rules of morality there will be a reward in terms of good feelings.

Evolutionary geneticists struggle with the concept of altruism; perhaps they have been too strongly influenced by Richard Dawkins' unfortunate choice of the term "selfish gene" [12]. Why should an individual risk their life to help someone else? Surely a genetic mutation would arise to eliminate the altruistic behavior and the mutant gene would be at an advantage in the gene pool. But this is not how genetics works. Behavior is controlled by large networks of genes not single genes. The networks are complex and have a high level of redundancy [13, 14]. Furthermore genes are used in different combinations to control different patterns of behavior [15]. Single deleterious mutations have little effect on behavior because of redundancy in the network. Combinations of deleterious mutations will interact synergistically to impair function and this is an important cause of disease and functional impairment.

But individuals with many deleterious mutations are at a disadvantage and the mutations are eliminated by natural selection. The selfish gene doesn't stand a chance. When we think about society and the place of human beings in society then altruism is of major evolutionary advantage. Individuals follow a moral code which includes altruistic behavior and everybody benefits. Individuals might have to risk their life but if they don't then they will suffer the feeling of guilt and might be ostracized by society. Cowards will not do well in terms of marriage and procreation. Risking one's life, however, is the exception. Altruism is more about helping others and getting pleasure as a result. There is mutual advantage to those who receive help in return. In evolutionary terms altruism is an advantage to the individual and to society.

HAPPINESS

Happiness has been the subject of scientific investigation in recent years [16]. In fact it is a scientific growth industry. Economists and politicians have realized, somewhat belatedly, that increasing gross domestic product (GDP) does not necessarily correlate with increasing happiness and increased well being. Parents have always claimed that what they want most of all is for their children to be happy; not necessarily wealthy and successful although that would be nice as well.

The measuring systems of this new science are simplicity itself. People are asked to score their sense of well being or sense of happiness on a scale of 1 to 10. Those who are comfortably off, in secure employment, married with a loving family have a high score. The unemployed, single and divorced have lower scores. In general happiness correlates positively with income up to a low multiple of the median or mean wage, but there is little rise thereafter. However the overall happiness level does not tend to rise as GDP rises over time. Interestingly this approach seems to work in that the scores are consistent between different studies. Rankings of activities in happiness order are also consistent within groups and over time. We tend to agree on the order when asked to rank sexual intercourse, eating a meal, drinking a glass of wine, reading a book, watching the television, going to the cinema, walking in the countryside, playing with our children, feeling sad, experiencing toothache etc.

According to decision theory we must be able to express anticipated pleasure and pain on a common single scale so that the numbers can be combined with scores of probability and an action initiated only if a threshold

value is exceeded [4]. Thus the fact that the happiness scale works is actually a prediction of decision theory (table).

BEAUTY

Beauty makes us happy. We like to walk in the country, we love beautiful gardens, we enjoy literature, poetry, music and drama. Indeed human beings, when not involved in the necessities of life, spend a great deal of time creating, admiring and enjoying beautiful things. We have an advanced abstract concept of beauty. Once again we need to think in terms of genetic networks underscoring the concept of beauty and neural networks learning to appreciate the beauty of the culture in which we grow and develop. But what is the evolutionary advantage, how and why has the concept of beauty evolved?

We have two sets of approximately 25,000 genes. One set from our mother and one from our father. But a small number of these genes are damaged by mutation. These mutant genes, called deleterious mutations, are effectively lost in functional terms. They either cannot be copied to produce proteins or they produce proteins that do not work or, even worse, proteins that interfere with other functioning proteins. The unit of information that the gene represents is lost. The mean number of deleterious mutations in the genome of adults is less than 10 and approximately one new deleterious mutation enters the genome per generation [14] (figure). Obviously there has to be some system of selection against the deleterious mutations or they would steadily increase in number. The deleterious mutations in germ cells in men and women are distributed randomly at meiosis and there will be a Poisson distribution of deleterious mutations in zygotes with a mean which is less than 10. Genes act in large complex networks. The networks are redundant (rule 4 above) and therefore the deleterious mutations will interact synergistically to impair network function. Zygotes with four or more deleterious mutations in any one network will abort. Thus the zygotes that develop will have a Poisson distribution which is shifted to the left and the mean of the ones that survive is one less than the mean of the original set (figure).

The zygotes that survive will form children and then adults. Those individuals at the left hand side of the distribution will have few deleterious mutations and no more than 1 or perhaps two per genetic network. The networks will function adequately. These individuals will do well on any tests of skill.

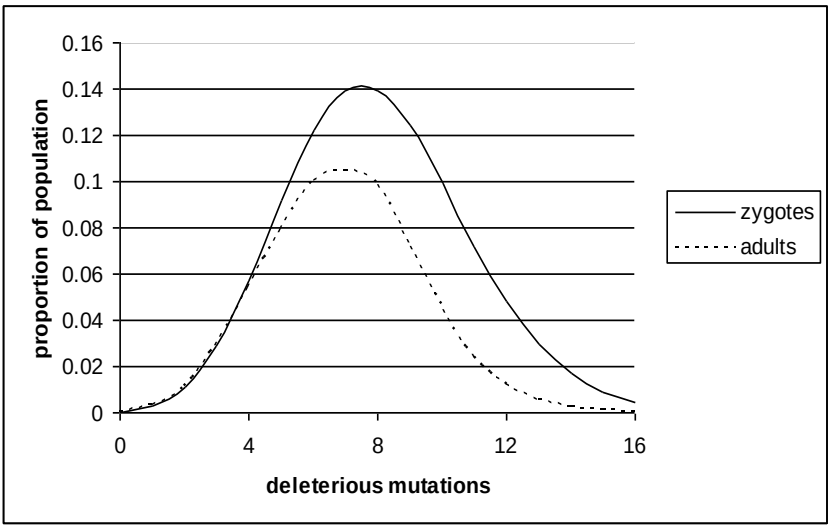


Figure. This graph shows a Poisson distribution of deleterious mutations in zygotes and a slightly skewed Poisson distribution in the zygotes that survive to reach adult life. The difference in mean between the two distributions is equal to the new mutation rate per generation. Adults towards the left hand side of the distribution have fewer deleterious mutations and there will be less impairment of their genetic networks. They will have higher scores on tests of skills including intelligence tests. They will be better at fighting disease and will have more symmetrical body forms. The genetic networks with no deleterious mutations will perform not just adequately but beautifully.

They will score well on IQ tests, they will do well at school, they will able to fight disease, they will have a symmetrical body form, they will get good jobs, they will marry and have children. Those at the other end of the distribution will be at a disadvantage. They will have several deleterious mutations in some of their systems. They will do less well on IQ tests, they will be more susceptible to disease, they will be less symmetrical in body form and they will do less well in life [17 – 20]. There is a clear and unequivocal genetic advantage if we can recognize those at the left side of the distribution and choose our sexual partners accordingly [19, 20].

In polygamous societies most women mate with just a few men and this is a powerful way of reducing the population load of deleterious genetic mutations. But even in monogamous mating systems selection based on the perceived level of deleterious mutations will bring down the overall population load [19]. Furthermore there is an advantage for those who can choose in the left half of the curve in that their progeny will be healthier. Intelligence, good

health and a symmetrical body form are easily recognized and do influence sexual attraction. We also recognize performance on a whole range of other skills such as sense of humor, kindness, altruism and morality which will also influence our choice of mate. Of course in a monogamous system the choice is not one way and those with the most gifts will tend to choose each other.

There is, however, a difficulty, if choice is based on function. In a redundant system there is little or no significant difference in function between 0 and 1 deleterious mutation. Indeed even with 2 deleterious mutations the system works well, it is only 3 that risks failure. You cannot choose the very best if choice is based on function in a redundant system; but what if the same genes are used in different combinations to produce beautiful ornaments such as the peacock's tail. Animals with an advanced concept of beauty will then be able to recognize the perfect pattern with no deleterious mutations. Humans do not have decorative tails but a symmetrical body form is a potent sexual attraction without an obvious functional advantage. In addition we have a whole range of creative abilities which are of limited functional value in the fight for survival but they seem to be of value in finding a mate.

Beautiful things have high information content. You can demonstrate this by photographing a painting and then gradually changing the pixels at random. As uncertainty is introduced and the information content is reduced the beauty fades. We seek sexual partners who are beautiful and can create beauty around us. If successful we are rewarded with pleasure and with happiness. The reward is delivered by our discerning non conscious mind to our conscious mind. The deep evolutionary advantage is elimination of deleterious mutations and preservation of our genome.

RELIGION

Religion is species specific behavior. All human societies, both primitive and modern, of which we are aware, develop and practice some form of religion. Like language and morality it is an emergent property of a human social group. Evolutionary biologists are forced to accept that in some sense there are genes for religion and that we all possess a religious deep structure, analogous to that for language.

The neural networks of the developing child are then attuned to recognize and learn the religion of the society in which the child grows up. Religion is concerned with altruism, morality, ritual, ceremony and beauty. Religious observance leads to pleasure. In so far as it is glue for society it has

evolutionary benefits. The individuals within the society are more likely to survive and procreate.

The rules of morality are written in the non-conscious part of our brains. We are conscious of the rules, at least in part, but we cannot re-write them by conscious effort. We also have a concept of an absolute right and wrong. This is not the view of any one person but the collective view of everybody in society. It is an emergent property of the group. Just as there is an English language and it is possible to say that the utterance of a particular individual is not grammatical. In the same way there is an absolute morality which allows us to declare the behavior of any individual as unacceptable. But where does the authority lie for deciding what is right and what is wrong? It is obvious that a functioning society requires such an authority. It cannot be a democratically elected government. That would be equivalent to our conscious minds being in control of the rules of morality. It isn't the judiciary or the legislature, both under conscious control. It isn't the modern monarchy. It has been the church, but that seems to be under threat. However an absolute morality will emerge bubbling up from the non conscious parts of our brains to impose its authority. If it isn't the church there is a danger that it is the mob. There is a lot to be said for organized religion with the collective wisdom of the ages setting standards in society. You do not need to believe every word in the bible to see that is a potential force for good in the world.

Religion is also about beauty and beauty is one of the deepest and most important concepts in biology. Few communities can afford to house everybody in palatial style but the community can build a beautiful church for collective worship. The church can be richly adorned and provide a place for those special occasions such as weddings and funerals. The community, through the church, can fund musicians and artists. The community can also fund schools and Universities. Secular governments try to copy some of these things but they never quite succeed. They are too concerned with function; beauty is much more important. I believe we can study religion scientifically. Enquiring into its neurological base and working out how it contributes to society and to our emotional well being. I would not pose the question: "Is there a god?". Because I do not know what it means. But we do need to know and understand the word of god. The absolute authority determining what is right and what is wrong. We will learn a lot about our collective non conscious neurology in the process and might find ways of enhancing our happiness and well being.

I hope that one day there will be world government. But we would need a common language, a common morality and a common religion. It is clearly feasible because we have a common deep structure for all three.

PSYCHOLOGICAL MEDICINE

I spent my professional career as a consultant histopathologist. This involved diagnosing disease on the basis of structural changes as seen down the microscope. Much of our understanding of disease is based on these structural changes and they underpin the way disease is classified. This applies to all body systems but not to psychiatry. I tell the medical students, not entirely tongue in cheek, that is why psychiatry never seems to make any progress. Psychological disease is classified in terms of common groups or sets of symptoms [21]. But we are rarely in a position to explain the underlying cause. However I believe that thinking in terms of decision theory and the interplay between the conscious and the non conscious part of the brain can help to formulate ideas. And if we combine those ideas with the science of genomics and proteomics real progress will follow.

Our mental well being depends on getting the balance right between the goals followed by our conscious self and the rewards and punishments delivered by our non conscious self. For instance there has to be a balance between selfish and altruistic actions. If we are too selfish then we will suffer feelings of guilt. If we are too altruistic we will fail to satisfy our selfish cravings. There is a balance that maximizes expected value which we need to find. My impression is that many people fail to get the balance right and suffer as a result.

There are optimum periods in childhood for learning; that includes language, the rules of morality and the development of immunity. The lack of a facilitating environment and traumatic experiences in childhood will inevitably affect mental health in adult life. Whether or not the developmental programs can re-written in adult life is still not clear. But talking therapy, cognitive therapy and behavioral therapy are all aimed at that goal.

One condition that can be understood in terms of the ideas presented in this chapter is post traumatic stress disorder. A young man grows up in a time of peace and learns a benign moral code. He then goes to war and is expected to kill his fellow man – the enemy. He then returns to his former life. Surely memories of his experiences in war will not go unpunished by his own moral

code. He will suffer anxiety and pain when he remembers actions that are unacceptable in peacetime.

It is a curious fact that men have the higher rate of mortality but women suffer more morbidity. In middle life women suffer a range of conditions many of which are caused by autoimmune mechanisms. They rarely kill but they can cripple; diseases such as rheumatoid arthritis, autoimmune thyroid disease, and multiple sclerosis. The decision theory model explains how those who minimize their risk of infection will increase their risk of autoimmune disease by producing antibodies that cross react with host tissue. Men follow the opposite course which puts them at risk of infection which does kill rather than cripple. Women are also more prone to mental health problems such as anorexia nervosa, chronic fatigue syndrome, irritable bowel syndrome and depression. This leads me to propose the following testable hypothesis: auto-antibodies arising because of molecular mimicry between gut microbes and brain tissue cause the aforementioned mental health problems which occur mainly in females [22].The antibodies cross from the blood to the hypothalamus where the blood brain barrier is deficient and react with receptors on the surface of neurons in the hypothalamus. The centers of emotion in the adjacent limbic system are then stimulated or repressed. In the case of irritable bowel syndrome the autoantibody target could well be the enteric nervous system.

DEPRESSION

If one observes a severely depressed patient it is not obvious what deep biological purpose is being served. Depression, however, is a stereotyped response. It overlaps with grief and extreme sadness. These states are specified by genetic networks which have been highly conserved in evolution. There must be a purpose.

A man is considering buying his son a car. There are all sorts of factors to take into consideration. There is the cost in money of the car and the cost of fuel, to set against the pleasure and convenience of owning the car. But there is also a risk that the young man could be injured or even killed. The risk of death might be low but the pain if death occurred would be high. The equation therefore would have a very high cost of the death of a son multiplied by a very low probability of the event occurring. The expected cost is low and therefore it probably would not greatly influence the decision. But somewhere in the non conscious brain is a very high cost associated with the death of the

young man and if he were to die then the cost would have to be borne. Thus there has to be a genetic system for causing grief and extreme sadness or the whole process on which conscious decisions are made would be rendered meaningless. Grief and sadness must be borne when a loved one dies and it is best not to interfere. Sometimes, however, grief can be prolonged and excessive and lead to severe depression which is a disease state. At other times severe depression can arise for no obvious reason. The question we need to ask is what pathological process could have switched on the genetic networks and proteomic networks which have caused severe depression. One possibility is auto-antibodies as indicated above.

PSYCHOSIS

Schizophrenia and manic depressive psychosis are relatively common psychiatric conditions each with a lifetime prevalence of approximately 1 in 100 [21]. They have characteristic symptoms which allow the diagnoses to be made with a reasonable degree of confidence. The conditions run in families and there is an underlying genetic etiology although the precise genes involved have not yet been identified.

The most likely genetic mechanism for both conditions is a combination of three deleterious mutations in a large genetic network of several thousand genes [14]. This fits well with the observed frequency in the general population and the risk of disease in close relatives. The degree of concordance in monozygotic twins is 50% and so possession of the three mutant genes alone is not enough. There are also environmental influences.

Genome wide association studies which measure a large number of common polymorphisms in the genome have failed to identify any genes with a strong effect in schizophrenia and in manic depressive psychosis [23, 24]. There are, however, some genes involved in immunity that have a small effect. This suggests that infection might have a role as a precipitating agent but it does not explain the underlying genetic cause.

The next stage in the genetic investigation will be whole genome sequencing. This should succeed in identifying the key deleterious mutations but the process will not be easy or cheap. The precise combination of three deleterious mutations will be different in different patients although there will be overlap in families. Once the genetic network is identified it should be possible to understand the underlying process of disease and hopefully interfere with the agents that precipitate exacerbations.

DISCUSSION

Consciousness and vitality are emergent properties of complex systems [3]. We cannot understand these concepts fully until we have worked out every detail of the complex biological systems from which they emerge. The science of molecular biology, genomics and proteomics in particular, has already provided deep insights into the nature of vitality. The same will apply to the nature of consciousness as we gather data on the human genome and proteome in health and disease. It will be a major task and an expensive task to gather the data but it will be eminently worthwhile.

There are, however, broad principles that apply to complex information processing systems and these can help in seeing the broad picture as the data emerges [2]. Conscious animals make decisions in an uncertain world and they must in some sense integrate *a priori* probability, evidence and the expected benefits and costs of right and wrong decisions [4, 5, 9]. Maximizing expected value for a conscious animal is seeking pleasure and avoiding pain. These are subjective sensations delivered to the conscious mind by the non conscious parts of the brain and body. The amount and complexity of information processing is not greater in conscious decision making than in subconscious or non conscious decision making. This is best illustrated by the working of the immune system [7]; our thought processes do not compare in complexity with the task of building and maintaining the immune armory. Consciousness, however, gives us the chance for new solutions to hitherto unforeseen hazards; solutions that others can copy and can then be passed down the generations. The acquisition of language by a child, the development of a moral code, the proclivity to religion and the appreciation of beauty are determined by neural networks which learn from the environment. The process is not under direct conscious control. Our conscious mind, however, uses these skills and attributes and is in turn controlled by them.

The evolutionary value of language to a social animal is obvious; but so to is the value of a moral code and a religious propensity. Human beings living in society are more likely to survive and reproduce. The benefits of society outweigh any costs of altruism to individual survival. The conscious mind, however, is not concerned primarily with survival and reproduction but with maximizing pleasure that comes with social intercourse, altruistic behavior, religion and the appreciation of beautiful things.

Parents want their children to be happy, or so they claim. Politicians and economists have realized that there is more to life than money and that there might be votes in happiness [16]. The decision theory model sets happiness as

the strategic goal. Happiness is a scientific growth industry. A curious aspect of the success of this field is that simple systems of measurement give consistent results. A wide range of emotions, sensations and experiences can be expressed on a simple one dimensional scale. The decision theory model predicts this finding as maximizing expected value presupposes that benefits and costs can be expressed in the same units on a common scale (table).

A great deal of unhappiness is caused by mental ill health. This can take many forms and there are undoubtedly many causes. A conflict between conscious goals and subconscious moral goals is likely to be responsible for a proportion of this discomfort. A possible role for autoimmune processes in young and middle aged females arises from the analysis presented above and if confirmed would point to a new approach to tackling mental disease [22]. Autoimmune disease is mainly a consequence of infection and low dose early mucosal exposure to micro-organisms is, according to information theory, the way to minimize this risk [2].

A small number of heterozygous deleterious mutations interact synergistically to impair the function of large complex genetic networks (figure) and contribute to the etiology of schizophrenia and manic depressive psychosis [14]. Genome sequencing will allow us to identify the networks and this will give insight into the nature of these conditions. Environmental factors are also important in precipitating psychosis. If we understand the genetics there is a good chance of identifying and modifying the environmental factors. The last big question in biology will be answered as we slowly unravel the molecular basis of complex systems. This insight into the nature of consciousness, however, will not explain away altruism, morality, beauty and religion. Instead it will enhance our understanding and appreciation of that which makes life worthwhile.

REFERENCES

[1] Lewin B. Genes IX. London: Jones and Bartlett Publishers; 2008.

[2] Morris JA. Information theory: *a guide in understanding health and disease*. www.lulu.com 2009.

[3] Morris JA. The conscious mind and its emergent properties; an analysis based on decision theory. *Medical Hypotheses* 2011; 77: 253 – 257.

[4] Green DM, Swets JA. *Signal detection theory and psychophysics.* London: John Wiley and Sons; 1966.

[5] Morris JA. Information: theoretical and applied. In: Progress in Pathology, Volume 5: editors Kirkham N and Lemoine NR; Greenwich *Medical Media Ltd.* 2001.
[6] Kahneman D. Thinking, fast and slow. *Allen Lane, Penguin Group*; 2011.
[7] Roitt I. Essential Immunology, 8th edition. Oxford: *Blackwell Scientific Publications*; 1994.
[8] Tannock GW. New perspectives of the gut microbiota: implications for future research. Gastroenterol Clin North Am 2005; 34: 361 – 83.
[9] Morris JA. Autoimmunity: a decision theory model. *J. Clin. Pathol.* 1987; 40: 210 – 215.
[10] Morris JA, Harrison LM. Hypothesis: Increased male mortality caused by infection is due to a decrease in heterozygous loci as a result of a single X chromosome. *Medical Hypotheses* 2009; 72: 322 -324.
[11] Chomsky N. Reflections on language. London: *Temple Smith*; 1976.
[12] Dawkins R. *The Selfish Gene*. Oxford: Oxford University Press; 1976.
[13] Morris JA. Information Theory: a guide in the investigation of disease. *Journal of Bioscience* 2001; 26; 15 – 23.
[14] Morris JA. Synergistic interaction of heterozygous deletions impairs performance and confers susceptibility to disease at all ages. *Medical Hypotheses* 2005; 65: 483 – 493.
[15] Morris JA, Morris RD. The conservation of redundancy in genetic systems: effects of sexual and asexual reproduction. *Journal of Bioscience* 2003; 28: 671 – 681.
[16] Layard R. Happiness: lessons from a new science. *Penguin*; 2005.
[17] Moller AP, Thornhill R. Bilateral symmetry and sexual selection: a metanalysis. *Am. Nat.* 1998; 151: 174 – 192.
[18] Moller AP, Mousseau TA. Mutation and sexual selection; a test using barn owls from Chernobyl. *Evolution* 2003; 57: 2139 – 2146.
[19] Morris JA, Morris KL, Morris RD. The mathematical basis of sexual attraction. *Medical Hypotheses* 2002; 59: 475 – 481.
[20] Morris RD, Morris JA. Sexual selection, redundancy and survival of the most beautiful. *Journal of Bioscience* 2004; 29: 359 – 366.
[21] White P, Clare AW. Psychological Medicine. In: Kumar and Clarke, editors. *Clinical Medicine. Elsevier*: 2005, p. 1273 – 1315.
[22] Acres M J, Heath J J, Morris JA. Anorexia nervosa, autoimmunity and the hygiene hypothesis. *Medical Hypotheses* 2012; 78: 772 – 775.
[23] Shi J et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; 460: 753 – 757.

[24] Ku CS, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: where are we now? *Journal of Human Genetics* 2010; 55: 195 – 206.

In: Philosophy of Mind
Editors: R. J. Jenkins, W. E. Sullivan
ISBN: 978-1-62257-215-1

Chapter 5

MIND IS CULTURALLY CONSTRAINED, NOT CULTURALLY SHAPED

Alessandro Antonietti and Paola Iannello
Department of Psychology
Catholic University of the Sacred Heart, Milan, Italy

ABSTRACT

The claim that mind is shaped by culture – as nowadays assumed by many authors and theoretical perspectives in philosophy, psychology and social sciences – is critically discussed. Firstly, three positions concerning the contribution made by culture to the alleged construction of the mind are highlighted: (i) culture contributes to the construction of the mind by offering *opportunities* that allow the endogenous psychological resources to be made explicit; (ii) culture contributes to the construction of the mind by producing *influences* on it; (iii) culture is the *condition* to construct the mind. The last position leads to problematic consequences which are contested. If mental experience arises by assimilating the dominant cultural framework, where do questions about the cultural framework itself come from? People sometimes realise that there is something which fails to fit the shared cultural framework. On the opposite, a cultural framework is accepted because it is realised that it gives an adequate explanation of aspects of reality that to a certain extent had already been perceived outside the framework itself. Finally, radical culturalism fails to give reason of why and how changes in the existing framework occur: if mind is shaped by current culture, where do original ideas come from?

These arguments lead us to concede that there are mental experiences not mediated by culture which are the source of debating and innovation. In conclusion, the culturalist perspective reminds us that mind is culturally constrained but can not induce us to believe that it is merely a cultural construction lacking of a genuine psychological status.

1. A "NEUROBIOLOGICAL" MIND?

According to a popular, even though questionable, view, psychologists are mainly asked to give reasons of atypical behaviours. A relevant example of a situation needing for a psychological understanding is the following. Some people report intense experiences labelled as "religious": they claim to have spoken directly to God or, less presumptuously, to have been bathed in a divine light illuminating everything, in such a way that an ultimate truth – beyond the capabilities of ordinary people – has revealed itself to them, leading them to perceive the intrinsic beauty of the world. These mystical experiences alter the personality of the person, who begin to be interested in spiritual and ethical problems, claim to see hidden and symbolic meanings in ordinary events and become hypergraphic (writing diaries, essays etc.). Some of them undergo true conversion. These subjects, who generally appear self-centred and arrogant, report of having touched the ultimate ground of reality, and experience a sense of union with the universe (Stace, 1960). In conversation they tend to quibble and be argumentative. This is the case, for instance, of Paul reported by Ramachandran and Blakeslee (1998), who generally emerges as a mystic, appears arrogant and presumptuous, claims to have received an illumination which allows him to understand the profound sense of reality and to feel a sort of mystical union with everything, spending his time on the compilation of huge volumes in which he illustrates his ideas and experiences.

In which sense is an explanation of behaviours like those reported above conceived as "psychological"? Usually, in the sense that such an explanation makes reference to "something" which occurs in people's mind. But what is the status of the mental entities, which are assumed to be the cause of people's behaviour?

Nowadays many scholars claim that the true nature of the alleged mental entities is constituted by the underlying neural mechanisms (Antonietti, 2008). For instance, the abnormal manifestations of religiosity mentioned before usually occur in conjunction with seizures experiences linked to temporal lobe

epilepsy (Devinski, 2003; Saver and Rabin, 1997). This led to maintain that the cause of those pathological behaviours is located in the temporal lobe, more precisely in possible lesions or alterations of the neural processes which result in the misfunctioning of that brain structure. In this field numerous studies concluded that during mystical or prayer states increased the activity in specific brain areas, exploring the neural underpinnings of this spiritual experiences (Beauregard and Paquette, 2006; Newberg, Pourdehnad, Alavi and d'Aquili, 2003). But, to what extent do these neurobiological mechanisms can be conceived as an explanation of the aberrant form of religiousness in question?

Nobody can deny that the work of the mind is accompanied by the work of the brain. In general terms, we can say that each mental experience corresponds to a biological process. Mind-brain correspondences are not surprising if the human being is conceived as a psycho-biological unit undergoing mental experiences with its whole body. "Non ex quacumque virtute quaevis actio procedit", as the Scholastic maxim goes (Thomas Aquinas, *Summa contra Gentiles*, cap. 2; see also *Quaestiones Disputatae cum Quodlibetis*, IV, q. 2, a. II, p. 508 of the Leonian edition): our bodies are not an indistinct biological plasma. Like other organic beings, the body is a structure highly differentiated internally whose functioning is highly organised. It is therefore obvious that the counterparts of an articulated and finely differentiated mental life are specific biological processes. It would be surprising if our mental acts were accompanied by biological activities without any specificity or regularity. Since man is a psycho-biological unit, I am not surprised to learn that when a person performs a certain type of behaviour or lives certain types of mental experiences, particular structures, specific forms of functioning and so on are activated in his/her nervous system. Such a parallelism is exactly what I expect. Thus, the fact that, when someone has a mystical experience or develops a "religious" personality, something peculiar occurs in his/her temporal lobe is quite natural. The question is: is what occurs in the temporal lobe the explanation of the mystical experience/religious personality (Antonietti, 2010b)?

What neurosciences probe is a concomitance of the mental experience M (mystical experience) and the neural activation N (temporal lobe). Such a correspondence can be conceived in four different ways:

1 co-occurrence without implication: M and N sometimes (perhaps even very often) occur together, though sometimes M occurs but not N or N but not M;

2 implication of the mental experience (M) on the part of the neurobiological process/structure (N): each time N occurs, M also occurs. M may however also occur without N;
3 implication of the neurobiological process/structure (N) on the part of the mental experience (M): each time M occurs, N also occurs. N may however also occur without M;
4 reciprocal implication: N always and only occurs together with M and vice versa.

The example of the temporal lobe syndrome certainly is not one of the second or fourth type, since not everyone who has epileptic focus in the temporal lobe shows "religious" attitudes (this form is documented in 38% of cases; in the other cases the subject addresses his/her interests towards art, philosophy, etc.). It could be one of the third type, if it could be ascertained that there is an elective involvement of the temporal lobe in all those who show similar sensitivities and "religious" behaviours; otherwise it would be one of the first type.

It is important to establish what type the correspondence in question belongs to, since different conclusions may be drawn (Antonietti, 2010a, 2011). For example, there is a difference between a mental act which is performed only in correspondence of a certain neurobiological event (which is therefore necessary for this act to take place, though it does not exhaust its function in sustaining the act, since it could take place even without this act being produced) and the case of a neurobiological event which takes place only in correspondence of a certain act (which might also take place concurrently with other neural processes and thus does not appear to have a specific biological counterpart).

It is worth noting, furthermore, that the correspondences in question lend themselves to different interpretations. One could be led to believe, if the correspondence is of the fourth type, that a certain structure of the brain is the “module” of the parallel mental act, that is, the element which contains everything required for this act to develop. But that structure might simply perform the role of initiator, catalyser, co-ordinator, controller, modulator, specifier or key element of a neural process which involves other structures and which only in its entirety is the actual biological counterpart of that act. Even wider is the range of interpretations possible if the correspondence is of the first, second or third type. Indeed it should not be forgotten that the activity of other neural structures is always present, even though in a certain circumstance one structure is found to be involved more than others. In any

case the fact that the structure N is mainly activated in concomitance with the mental act M or that the structure N is not present in the individual who is unable to perform the act M does not mean that N is the place, the cause, the determining factor of M. For example, the fact that a plant dies if substance Sx is lacking from the soil does not mean that the plant lives only thanks to Sx.

In addition, what can these correspondences tell us? In general they are used as evidence supporting distinctions between systems, components, functions, mental attitudes, often performed with the intention of showing their specificity/selectivity.

However, we can say that often there was no need of neurobiological data to know that given a certain psychological experience has its own distinctive character (Antonietti and Iannello, 2011). For instance, what can we learn from the temporal lobe syndrome phenomenon? We can conclude that the recurrence of critical experiences can change the personality, prompting to develop specific interests, attitudes and behaviours. But both the peculiarity of mystical feelings and the psychological characteristics of the "religious zealot" were already known because of psychological analysis, as well as the diminishing border between self and reality and spatial disorientation that appear during meditation which accompanies – according to data of Newberg and Aquili (1998) – the reduced activity of the upper posterior parietal lobe. Indeed, without prior psychological analysis, what would the neuroscientist relate the brain processes that he is investigating to? Sometimes, however, the psychological distinction that neurobiological data intend to prove it is not obvious.

Though the correspondences in question cannot prove pre-existent psychological distinctions, they can however suggest new ones. In the case of epilepsy of the temporal lobe, the revealed concomitance would prove that a "religious" experience such as that of these patients is special with respect to everyday life because the brain is in a special state when it occurs. But this type of experience was already known to be special, independently of neurobiological data, since thoughts, interests, attitudes and behaviours testified abundantly in favour of its specialness.

What is more, as we said before, it comes as no surprise to that psycho-biological unit which is man to know that a special psychological experience is accompanied by a special neural state. It could instead be interesting to discover, by means of further investigations, that – the example is completely fictitious the patients in question systematically show, during an epileptic seizure, either a N1 neurobiological activation or a N2 activation. This could be the evidence of two different forms of "contact" with the divine that had not

been revealed at a psychological level. A more detailed reconstruction of what patients experience or think could thus lead to identification of a contact of the M1 kind characterised by the sensation of understanding the mystery of reality (a predominantly intellective contact) and a contact of the M2 kind characterised by the sensation of serenity and beauty (a predominantly emotive-aesthetic contact). The differentiations suggested by the biological datum would need however to be proved on the psychological level by introspective accounts, analysis of attitudes, assessment of facial expressions, study of linguistic expressions and so on.

2. A "CULTURAL" MIND?

If neural mechanisms can not per se constitute the explanation of the psychological causes of behaviour, can such a role be played by culture? How the above-mentioned case of Paul can be interpreted from a culturalist viewpoint, another theoretical perspective which is shared nowadays by many scholars? We could say that Paul periodically goes through new and disorientating experiences (during epileptic seizures). Paul tries to interpret these experiences by making reference to a conceptual system which in his cultured western mind he labels as "spiritual", "mystical" and so on. Paul tries to understand better: he takes this "religious" definition of what is happening to him (in another context he could have made use of an interpretative framework of a "mythological", "magical", "science-fiction", etc.) and goes deeper into it; he reads, writes, perhaps gets into contact with a religious group. These frequentations – literary and personal – lead him to assimilate a certain language, encourage him to follow particular practices which become part of his daily behaviour. Gradually Paul absorbs the "canons" of a particular religious micro-culture and constructs a new identity for himself. However the old self continues to live within Paul and it is this self that instils in him the idea that perhaps there might be something "strange" in what is happening to him; this "normal" self leads him to seek medical advice, go for clinical tests and so on. In private and within the circle of his religious companions, Paul is the mystic who is in touch with divinity, who wants to communicate to others what he has understood of life and the world and so on; in the neurologist's consulting room, Paul is the patient who talks about the "symptoms" of his disorder.

Also the claims of the culturalist approach merit to be thoroughly examined. Drawing ideas from the Soviet historical-cultural school, someone

maintains that the most significant elements for psychological dynamics are the instruments - particularly the social instruments - used by man. Man's relationship with reality, it is claimed, is fundamentally mediated by instruments and these shape the structure of the mind. Of these instruments, a special role is attributed to language, which has not only a communicative function, but also the role of organising and structuring the mind. Language, in turn, is not an absolute element, but still a coherent element of the cultural structure of which it is part. Since the same individual finds him/herself participating in more than one linguistic-cultural world, different "minds" are thought to develop within the same individual, and these minds are activated differently according to situations, circumstances, roles and objectives of the moment. These are the instances claimed by post-modern psychology. Modern psychology is thought to be based on the assumption of the possibility of having a common object of investigation and the possibility of searching for general laws of the mind. Instead, post-modern psychology assumes that the object of psychology is an historical construction, conveyed by the language, influenced by the historical-social context and therefore no universal laws exist because the mind is always historical. Furthermore, the psychological structure is not unitary. Man is decentralised and the subject is dissolved in the linguistic structures and in the sets of relationships in which it is involved; the self breaks up into fragments and multiple selves exist within the individual.

Along these lines, Harré and Gillet (1994) claimed that psychology must study how I narrate myself, how others narrate me; the mind is a system that controls the activity of people who interact in a specific spatial-temporal and socio-cultural context; the mind is the point of encounter of structuring influences coming from the context. These positions are supported by philosophical positions according to which multiple versions-interpretations-narrations co-exist and the mind is a constructed and reconstructed text, a space with multiple entrances and intersecting itineraries.

This is the general framework within we can collocate different positions sharing the same conviction that cultural-contextual aspects structure the mind. As in the case of the attempts to give reason of mental experience exclusively in neurobiological terms, also the attempts to give reason of them in cultural term need some conceptual distinctions.

Within the cultural framework we can identify three positions concerning the contribution made by the cultural context to the construction of the person.

One possibility is to claim that culture contributes to the construction of the person by offering *opportunities* that allow the endogenous resources of the individual to be made explicit. In this perspective the contribution of

culture is relatively extrinsic, since mental activity is still considered as mainly generated from within the individual. Culture is seen as a factor of facilitation-hindrance, acceleration-deceleration, and so on, but not as a determining factor. Culture concerns the variables which allow or prevent the disclosing of pre-existing potentials. This is the perspective of cultural conditioning.

A second possibility is to claim that culture contributes to the construction of the person by producing *influences*. Culture is seen here as a factor external to the individual which affects his/her development, directing its course. In this view, culture is seen as a complex of collectively shared ideas which in the mind of each individual are "translated" into personal models of interpretation of the specific situations encountered in real life and in prototypes of actions. Through these subjective representations, the collective ideas guide the behaviour of individuals. In this position, therefore, culture influences the process of construction of the person, leading him/her to develop a particular view of reality, to share certain social norms and values, to "read" his/her own experiences with particular emphasis, and so on. This is the perspective of cross-cultural psychology (Berry et al., 1992).

Lastly, it is possible to consider culture as *condition*, a constituent element of the mind of the individual. You cannot become a person without the contribution of culture. In this perspective – which is that of cultural psychology (Cole, 1996) – culture represents at the same time both the scenario in which mental activity is performed and the instruments with which it is produced. Culture provides the keys for interpreting reality, or simply the keys for accessing reality, since reality is intrinsically a cultural construction. To give sense to the experience, the individual – so it is claimed in this viewpoint – has to make use of frameworks of interpretation and these are offered by culture. Since the self is also a cultural construction, cultural membership determines the way with which the individual lives, represents him/herself, gives value and meaning to his/her own existence. If this view is adopted, the individual has to develop multiple identities (transactional selves) according to the contexts in which s/he has to interact.

The radical culturalist position constitutes a considerable challenge. Indeed, while in the first two positions the incidence of culture involves peripheral dimensions of the mind, in the last position culture affects the very core of the mind. Taking this perspective to the extreme leads to certain problematic consequences.

First of all, if the individual cannot be other than how culture has constructed him/her, where can there be space for freedom? It then becomes difficult to identify values, ideas, principles, views, interpretations and so forth

on which to converge and around which to create consensus (rationally, empathetically, etc.) since everyone is "imprisoned" in the cultural frameworks with which they have grown and from this there is no escape. Lastly, the individual is barred from the possibility of personalising the sense of the experiences because these must always be interpreted through cultural frameworks and these are always socially shared frameworks. Thus, since people have to apply to the majority, they do not capture originality.

If someone interprets reality by assimilating the cultural framework that applies in his/her environment, where do questions about the cultural framework itself come from? It should provide a suitable – though not necessarily exhaustive – interpretation of the world, yet people sometimes realise that there is something "not right" and ask questions about aspects of reality which are not provided by the framework. So what are they comparing the framework to? There must be a perception of reality not mediated by the framework which appears discordant with the framework and leads to questions being asked.

But even before all this: what motivates people to acquire the framework? It cannot always be only by mere acquiescence or imposition. A framework is accepted because it is realised that it gives an adequate explanation of aspects of reality that to a certain extent had already been perceived. It is accepted because it gives a good account of what was obscurely pre-felt. In order for the framework to be shared, what it proposes has to "stick", "coagulate", make resound something that the individual already pre-understood.

When I see a hedge against the horizon at night, my perceiving that hedge certainly cannot fail to be mediated by the cultural artefact of Giacomo Leopardi's poem *L'infinito* [The infinity] which leads me to perceive meanings and resonances in my current experience that would not emerge if I had not assimilated the "poetic framework" of hedge. Nevertheless, the first time I read that poem I perceived it as important because it linked up within me to meanings which I had up till then represented to myself in an obscure way. The fact that now my perception of that object «che il guardo esclude» [which prevents me from seeing] and the expression of what it awakens in me can only be mediated by the framework of Leopardian lyric does not mean that at the basis there was not an original encounter between something which was in any case present to me and a linguistic-conceptual form which I have found nicely adapts itself to that something. If it is true that culture is a set of instruments that help me to read reality, the acceptance of the readings that these instruments propose must be based on a perception of reality which is not culturally mediated. For example Proust in the last volume of of *La*

recherché du temps perdu [The search of lost time] says that literature is a sort of telescope/microscope through which the reader can see within himself, interpret himself, reveal himself to himself. But a blind man would draw no benefit from a telescope or a microscope, just as these instruments would have no use if there were no stars to look at. Leaving behind the metaphor, cultural instruments presuppose a reality to be interpreted (as opposed to the idea that reality is all a cultural construction) and a pre-relation (that the instruments can expand, refine, modify and so on) of the observer with this reality.

On the other hand, what can lead to cultural changes which is intrinsic to the individual and not just the effect of external agents? If a person is unable to access experiences of meanings that do not fall within the current cultural frameworks, where do new or different frameworks come from? The individual is sometimes aware of the gap between his/her own perception of the world and the cultural forms which this perception tries to approach. This gap seems to be symptomatic of something that the cultural frameworks cannot grasp.

Furthermore, even if it was admitted that the intersubjective comparison and the acquisition of cultural instruments of a community was the determining factor in the construction of the representation that an individual has of him/herself, this would not be the same as saying that personal identity is a cultural product. In short: the image of the self is not the self. It might be that, within the context in which I collocate myself, I am perceived in a different way, and I might even perceive myself in a different way, or the way my mind works might even differ according to the situations, but this does not mean that within me there are several minds, since such differentiations would always be made within a unitary dimension. (Similar comments can concern this argument: since many different discourses are developed about the mind, then there are many minds. Multiplicity of discourses does not imply multiplicity of their referents. Many different discourses can be developed about the same, unique mind, each discourse highlighting aspects neglected by other discourses. This does not mean that each discourse "construes" a different mind).

Lastly, it seems that we cannot preclude a priori the possibility of making generalizations about the psychological dimension. Firstly, some form of generalization is still present in every discourse: even when I describe an individual in its concrete life environment, contextualising what happens in his mind, I usually employ abstract concepts involving generalizations. Maintaining that such a person is experiencing an intense spiritual experience of communion within a certain religious group implies general knowledge

(based on generalizations) such as "spiritual communion" or "charismatic" group. You can not avoid generalizing. Secondly, if humans are so defined is because they share common characteristics that qualify them as "human" and that set them apart from other living beings. To a certain level then there must be something general – and not just a list of the specific characteristics of each individual who composes the human race – in all men.

Rather, it is worth evaluating case by case at what extent it is possible to ignore the concreteness of the individual case and generalize the description of psychological characteristics of persons, i.e. to determine to what extent the generalization does not make neglect those features without which the psychological phenomenon no longer has meaning.

The culturalist perspective reminds us that man's coming into contact with the sense of things is mediated by cultural constraints as well as by biological limits. The mind always operates in a cultural "humus" and is nourished by this. A person is also constantly engaged in interacting with instruments, individuals, values, behaviours, discourses which incorporate cultural dimensions. But recognising that the mind does not think and feel in a vacuum of culture does not mean that it lacks autonomous powers.

Going back to the case of Paul seen from a culturalist viewpoint, it is not surprising that certain frequentations - social, cultural and so on – lead a person changing his interests, habits, values. But this does not mean that they graft a second identity onto the same individual. Despite the change and despite the contradictory coexistence of different aspects, it is still the same Paul - we would not speak of change and transformation, but of substitution and juxtaposition if the referent were not the same individual. It comes as no surprise that the human person is not "monolithic" but subject to mutations. The latin motto «frangar, non flectar» (I can be broken but never bent), however noble, is certainly not the most followed one and it is perhaps not even the most practical, though we would not go as far as to say that the best way – as in the motto «fluctuat nec mergitur» (He who rises with the wave is not swallowed by it) – is continuous and opportunistic variation. In this motivated changing, the role of culture can be significant. But this does not mean that culture shapes the person.

Why do temporal lobe epileptics focus on religion, painting, philosophy and so on and not on mathematics or circus arts? Because in their experiences during the seizures there is a particular range of meanings which appears to them. If they choose to look deeper into their experience and an adaptation to them of their own behaviour within a religious framework, it is because the experience contains something that certain cultural frameworks can clarify,

reinforce but not determined from the start. If Paul finds his own way in "conversion" to culture and to "religious" practices, it is because the elements of his culture were already inscribed, though in nuce, in the original experiences and for this reason Paul has perceived them as syntonic with respect to these experiences.

It is known that there are experiences which are particularly significant for the individual, though they are not expressed within cultural frameworks. There is a way of feeling and intuiting with reference to reality and existence which appears to be "primitive", in other words preceding the cultural forms with which individuals then represent it. It appears that there must be a non-cultural primum, which has meaning for the individual otherwise s/he never would capture the sense of what culture "frames"

For example, it is true that the feeling of guilt takes on different aspects in different cultures (Lillard, 1998) (in some cultures it relates to the physical entity of evil caused – the more I feel guilty, the greater the damage I caused – while in others cultures it refers to the intention – I feel more guilty about having voluntarily slapped a person than about having involuntarily broken an arm); however, feeling guilty per se is not produced by culture, it is prior to culture, indeed.

Culture modulates, orients, accentuates an emotional-affective disposition whose meaning is not given by the cultural setting in which it is inserted. Shared social norms, values, beliefs alone cannot be sufficient to explain how a certain feeling emerge in a person. Or, quoting another example, it is true that the temporal perception varies from culture to culture (Golia and Antonietti, 1992) (in some contexts the same temporal durations are assessed as longer than in other contexts), but the feeling of spending time is a primitive experience that habits, narratives or schemas can only modulate, but not create from nothing.

What is more, being influenced and shaped by a certain culture and receiving from it a certain imprinting does not mean that the individual cannot have access to sensitivity, views, values that are typical of other cultures, or universal. In culture, a person finds "structures" which offer means for refining and expressing his/her own experience of reality. These structures are however "filled" with meanings which the individual can access and which can be moulded to individual requirements and intuitions which have not been foreseen.

In the case of Paul, we could say that at a certain point of his life an event happens to him (happens to him, to his person as a whole, not to his brain, because this neurological deficit modifies him in all his being, in his

behaviour, attitudes, fears, thoughts and so on) which makes him perceiving particular meanings, possible "projections", opportunities. Paul's intentioning acts, accompanied by his brain affected by the epileptic focus, are structured in a stable way (for instance, they correspond to certain stable interests, and so on) that orients towards certain cultural instruments and social aggregations, which reinforce, expand, give better expression to those interests, so much that Paul's personality is bent in a certain direction, making him take on a particular image of himself and particular relational styles.

It is certainly true that every individual opens to reality according to his/her own personal perspective, but the subjectivity of the experience should not be identified in something special that is determined in the mind and which differs from individual to individual. The singularity of the mental experience – just like the discrepancy between it and the actual configuration of things – does not derive from the fact that each of us has in his/her head a different model or representation of reality, but from the particular way with which the meaning that lies in things is perceived by the individual. In reality there is an intrinsic meaning which can be recognised by more than one person (and taking the same attitude towards it). Experiences are different not because minds are different boxes in which distinct things are put, but because there is the possibility of interpreting the world in original ways. Reality is "the same" for all minds – and this makes intersubjective agreement possible – even though it lends itself to being perceived in different ways by different individuals. Intersubjective agreement is not constructed operatively starting from radically different experiences, but from the participation in a common sense inscribed in things.

3. A MIND BETWEEN NEUROBIOLOGY AND CULTURE

The two challenges that today psychology is faced with – that in the preceding paragraphs we have tried to examine critically with respect to their conceptual bases – can be perceived as an opportunity for the discipline to remind of instances that sometimes tend to be neglected, to redesign its object of investigation, and to identify new research procedures through an active comparison with the epistemological reasons which a neurobiological and a culturalist perspectives propose and support. The knowledge about what is happening in the nervous system during specific mental acts helps psychology not to forget that such acts are always rooted in a body. It emerges the limit of those conceptions according to which the mind is a realm of representations or

models of reality and of algorithms that find justification in their internal logic, regardless of the characteristics of the physical basis that it implements. Instead, it can be recognised the value of certain claims by the "evolutionary psychology" (Cummins and Cummins, 2002) – which aims at reconnecting biological processes within psychological theories or guidelines, variously referred to as «adaptive behavior», «autonomous agents», «behavior-based cognitive psychology», «bottom-up cognitive science» (Bakker, 2000) – that assumes that there are not infinitely many possible ways to implement a certain mental capacities (since from the biological point of view only a few possibilities are plausible) and in which agents are designed as systems with brain and body interacting with the environment. This body-mind rooting, both phylogenetic and ontogenetic, should, inter alia, help to overcome dichotomies that sometimes research undergoes (for example, unconsciousness-consciousness distinctions, explicit-implicit levels, automatic-controlled systems, etc.). If the mind is inherently connected to the body, intentional dynamics which characterise it innervate also what happens below the threshold of consciousness and that is usually placed on the level of the neurophysiological mechanisms: automatisms, subliminal perception, crepuscular states are already crossed by an orientation towards openness, even thug not fully deployed, to meaning.

On the other hand the cultural perspectives warn against the tendency to make undue generalizations, to develop greater attention towards the historical and contextual dimension of mental life, to avoid developing a solipsistic view of the psyche: what you build in mind does not depend solely on the individual, but incorporates elements that come from the social and cultural exchange; the individual mind fits into a temporal relational continuum that feeds it and by which it is enriched.

The real challenge for psychology is then to find a suitable way to consider what neurobiology and culturalism suggest, going beyond the simple juxtaposition of respective contributions. By only juxtaposing these levels an effective progress in the knowledge of the mind becomes impossible. It would be necessary also to go beyond the solution of the plurality of levels or reading keys, which would be equally unsatisfactory (Antonietti and Iannello, 2011). The attempt may be to show how the neurobiological and cultural data, once compared to mental acts, "reflect" a psychological meaning contributing to broaden the understanding of the way in which man interprets reality.

REFERENCES

Antonietti, A. (2008). Must psychologists be dualists? In: A. Antonietti, A. Morandi Corradini and E. J. Lowe (Eds.) *Psycho-physical dualism today. An interdisciplinary approach* (pp. 37-67). Lenham, MD: Lexington Books (a division of Rowan and Littlefield Publishers).

Antonietti, A. (2010a). Do neurobiological data help us to understand economic decisions better? *Journal of Economic Methodology*, 17, 207-218.

Antonietti, A. (2010b). Emerging mental phenomena: Implications for the psychological explanation. In: T. O'Connor and A. Corradini (Eds.), *Emergence in science and philosophy* (pp. 266-288). New York: Routledge.

Antonietti, A. (2011). What does neurobiological evidence tell us about psychological mechanisms underlying moral judgment? (pp. 283-298). In: J. J. Sanguineti, A. Acerbi and J. A. Lombo (Eds.), *Moral behaviour and free will. A neurobiological and philosophical approach*. Rome: IF Press.

Antonietti, A. and Iannello, P. (2011). Social sciences and neuroscience: A circular integration. *International Review of Economics*, 58, 307-317.

Bakker, B. (2000). The adaptive behavior approach to psychology, *Cognitive Processing*, 1 (1), 39-70

Beauregar, M. and Paquette, V. (2006). Neural correlates of a mystical experience in Carmelite nuns. *Neuroscience Letters*, 405, 186-190

Berry, J. W., Portinga, Y. H., Segall, M. H. and Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications*. Cambridge, MA: Cambridge University Press

Cole, M. (1996). *Cultural psychology*. Cambridge, MA: Cambridge: Harvard University Press

Cummins, D. and Cummins, R. (2002). Innate modules vs innate learning biases, *Cognitive Processing*, 3 (3-4), 19-30

Devinski, O. (2003). Psychiatric comorbidity in patients with epilepsy: implications for diagnosis and treatment. *Epilepsy Behavior* (Suppl. 4), S2-S10

Golia, B. and Antonietti, A. (1992). Production of short time intervals in African immigrants, *Perceptual and Motor Skills*, 74, 893-894

Harre, R. and G. Gillett. (1994). *The discursive mind.* Thousand Oaks, CA: Sage Publications

Lillard, A. S. (1998). Ethnopsychologies: Cultural variations in theory of mind. *Psychological Bulletin*, 123, 3-33

Newberg, A. B. and D'Aquili, E. G. (1998), The neuropsychology of spiritual experience, In: H. G. König (Ed.), *Handbook of religion and mental health* (pp. 75-94). San Diego, CA: Academic Press

Newberg, A. B., Purdehnad, M., Alavi, A., and D'Aquili, E. G. (2003). Cerebral blood flow during meditative prayer: preliminary findings and methodological issues. *Perceptual and Motor Skills*, 97, 625-630

Ramachandran, V. S. and Blakeslee, S. (1998). *Phantoms in the Brain: Probing the mysteries of the human mind*. New York: Morrow

Saver, J. and Rabin, J. (1997). The neural substrates of religious experience. *Journal of Neuropsychiatry and Clinical Neuroscience*, 9, 498-510

Stace, W. T. (1960). *Mysticism and philosophy*. New York: Macmillan

To be included in Nadya Gotsiridze-Columbus (Ed.)
Philosophy of mind
Nova Science Publishers, Hauppauge

In: Philosophy of Mind ISBN: 978-1-62257-215-1
Editors: R. J. Jenkins, W. E. Sullivan

Chapter 6

COMPUTATION IN MIND

Françoise Chatelin*
CERFACS and University of Toulouse, France

"Nous ne sommes que des nains perchés sur les épaules des géants; nous voyons ainsi plus loin et mieux qu'eux, mais nous ne serions rien s'ils ne nous portaient de toute leur hauteur." [1]

Bernard de Chartres (12th century)

Abstract

Like Philosophy, Mathematics deals with abstract ideas, i.e. immaterial objects which inhabit and work in the Mind. The chapter "Computation in Mind" proposes to use the power of computational mathematics to explore some of the universal ways by which each human mind builds its "imago mundi", its image of the world. The primary focus is put on epistemology and the use of mathematics is minimal, relegating the necessary technical details to an appendix.

The Chapter develops the viewpoint that Science and Mind are mirror images for each other which use specific calculations over three kinds of numbers. It presents some epistemological consequences of the lack of associativity or commutativity for the two basic operations which are $\times$ and $+$ when the calculations are performed over vectors or over matrices. The evolutive nature of the scientific logic is illustrated

*E-mail address:chatelin@cerfacs.fr

[1] We are but dwarfs perched on the shoulders of giants; thus we see further and better than they did, but we would not be anything if they did not carry us with their full height.

on several examples. In particular induction in computation suggests that any matrix ring can be usefully considered as a structure of macro-scalars.

Keywords: Mind, body, matter, mathematical computation, scalar, vector, matrix, multiplication, addition, commutativity, associativity,eigenvalues, singular values.

Introduction

A preliminary word of caution

The reader should take this Chapter for what it is: the bold venture of a mathematician into the neighbouring domain of Philosophy. From remote Antiquity until the Middle Ages, there existed no continental divide between these two regions of human thought: like Nicole Oresme in the 14th century, any thinker would happily wander from one region to the other, ignoring that he was trespassing immaterial borders which would only be established in future times.

The scientific assumption

A paradigm shift occurred during the late Renaissance period in Europe, which imposed the primacy of matter over thought as the correct way to decipher rationally the world. The denial of the role of mind in the human understanding of the world enabled the objective western science to blossom, filling the world with more and more inventive machines. Techno-science has gone a long way into the material direction, creating a mechanical layer insulating man from Nature. There are great benefits for the *material* well-being of many in the affluent countries: hard or repetitive labour is delegated to machines, chemical drugs help fight physical and psychological disorders, etc ...

These achievements are so impressive that they blind us to the other side of the coin: the *spiritual* well-being of man is lagging behind in oblivion. True to their materialistic assumption, many western scientists claim that their mechanistic model of evolution is the *only* one which can work in the physical world. But of course no mechanistic model of human development can be sustainable in the long run because of the physical limitations of the planet which are now felt globally (desertification of the land, decrease in biodiversity, pollution, to name a few examples). The grip of matter over the human mind is so strong that it forbids many of the best minds to realise that the iron rule imposed by the world's new financial order is far from being a necessity, a law of nature. The so-called "laws of economy" are a product of the human mind

with a degree of reality much fainter than the universal law of gravity... To assign money and economic growth as goals rather than view them as means for human evolution is to worship a post-modern reincarnation of the antique god known as the "Golden Calf". In addition it is unethical to present as mandatory the compliance to a financial system deviously crafted to bestow on the richest minority worldwide the wealth created by the poorest multitude.

This unsatisfactory state of affairs reveals some adverse consequences of the implicit scientific assumption about primacy of matter over mind. The scientific assumption has played a major role in the swift development of experimental sciences during the past three centuries. However there is today a discomforting awareness that **Life** itself is the main challenger of the basic assumption of science, see for example [Husserl (1970), Klein (2008)]. More and more phenomena are uncovered in life and social sciences which lie beyond the sole grasp of the matter-based approach provided by physics, chemistry and molecular biology. Collective cooperation (rather than the individual struggle and fight against all others which underlies neodarwinism in biology) appears to be a winning mode of survival on the planet [Axelrod (1984)].

The mind-body connection

The many black clouds hovering on a human future driven by technology alone suggest to reconsider the mind-body connection: the road which links the mind to the body in any human being can be travelled in *both* directions. The two directions have been exemplified in Ancient Greece by the opposition between Aristotle (*Physis*) and his former teacher Plato (*Timaeus*). The same opposition was reenacted in 1989 as a celebrated debate between two first rank scientists at the Collège de France, the materialist neurologist J. P. Changeux (brain $\Longrightarrow$ mind) and the idealist mathematician A. Connes (mind $\Longrightarrow$ body) [Changeux and Connes 1995]. The realist stance is a necessity for experimental scientists whereas the idealist attitude is common among mathematicians. To be successful a dedicated researcher, either an experimentalist or a theoretician, has to believe in the "reality" of the subject of his "quest for truth" necessarily taking place in a human intellect, even if the outcome has to be mediated by a massive apparatus for the former, or by computer, pen and paper only for the latter. But of course the mental reality of the experimentalist differs vastly from that of the theoretician: they both express in different tongues two valid – albeit limited – aspects of a larger reality.

The analytic tradition in philosophy tends to oppose the two views– realist versus idealist – of the world. Such a methodological separation is a temporary help but becomes limiting when held as absolute for too long. There is so much traffic between mind and body at all times that one can easily conceive of an uncountable infinity of intermediate views to reconcile the extreme ones characterised by *hyle* (matter) and *holos* (whole). These concepts have always been the two focal points which helped man to think dialectically the mind- body processing of information. This age-old dynamical approach to Life has survived in Eastern philosophies until modern times, whereas it was gradually replaced in the West by the scientific approach which resulted in a vast increase in the technological know-how. This was made possible by the remarkable *theoretical* development which took place for mathematics and physics in the 19th and 20th centuries. One cannot overestimate the epistemological importance of the rational light that these two advanced sciences concur to shed on our intellectual know-why for life on earth. Such a comprehension is mediated by *mathematical computation* in the mind, a mechanism which has so far received little attention from philosophers and scientists.

In this Chapter we shall review some principles which underscore the intellectual mechanism of "Computation in Mind".

Mind and Science

The scientific "imago mundi"

"Why do we see the world the way we do?". This must be one of the key questions repeatedly asked, at first implicitly, since of the emergence of *homo sapiens*. The scholastic version of this question was: "How do we construct our *imago mundi*?" Various answers have been proposed through millennia by religions and philosophies to the nagging questions of the mental construction of a working image of the world shared by a given social group.

The overwhelming variety of answers which have been developed in time and space on the planet comforts social scientists in their relativism: there cannot exist any universally valid answer. At the other end of the spectrum western science puts a strong claim to universality. This claim is based on another type of experimental evidence: the technological world (from cars to mobile phones and GPS) is based on universally valid scientific principles. These divergent opinions have nevertheless a common source in the human

mind! We shall concern ourselves below with the scientific construction inside the human mind of the contemporary imago mundi.

Science as the mirror of Mind

"What is reality?", this is an unfathomable question that haunts mankind since the origins lost in the mists of times. This primary question calls for the next interrogation: "Does reality exist out there or is it a mental construct?"

Modern science emerged in the 17th century by conceiving the totality of what is (the universe) as strictly divided in two radically different worlds: on the one hand, the *outside* world – representing "objective" reality, the primary domain of scientific investigation – on the other hand, the *inner* world – representing the "subjective" reality of each individual which lies beyond the reach of science. The early black-and-white picture has become fuzzier four centuries later. What looked like a waterproof wall between in and out is permeated by image processing. Some of the physical activity of the brain is within the reach of scientific investigation. We know more about the "how" in the mind. But what about the "why"?

Only the "why" can shed some light on the validity of the objective/subjective dichotomy posited by science. Such a dichotomy made no sense for Protagoras of Abdera (ca. 480-411 BC) who claimed the preeminence of the subjective in his famous statement: "Man is the measure of all things". Many centuries later, Poincaré arrived at the following synthesis between in and out based on human mind: "...la seule réalité objective, ce sont les rapports entre les choses d'où résulte l'harmonie universelle. Sans doute, ces rapports, cette harmonie ne sauraient être conçus en dehors d'un esprit qui les conçoit et qui les sent. Mais ils sont néanmoins objectifs parce qu'ils sont, deviendront, ou resteront communs à tous les êtres pensants."[2] [Poincaré (1905), p. 184].

The lesson of Protagoras applies to modern science even if some scientists tend to forget this lesson. The eternal objective/subjective dilemma was lucidly described in the 1930s by the mathematician T. Dantzig, a former student of Poincaré:

"The man of science will act *as if* this world were an absolute whole controlled by laws independent of his own thoughts or acts; but whenever he discovers a law of striking simplicity or one of sweeping universality, he will be wise to wonder what

[2] ... the only objective reality consists of relations between things from which stems universal harmony. Admittedly, such relations, such harmony cannot be conceived of in the absence of a mind to conceive of and feel them. However they are objective because they are, will be, or will remain shared by all thinking beings.

role his mind has played in the discovery, and whether the beautiful image he sees in the pool of eternity reveals the nature of this eternity, or is but a reflection of his own mind." in [Dantzig (1954), p. 233], italics in the original.

This quotation taken from Dantzig's book titled "*Number. The language of Science*", summarises in a poetic language the message of this Chapter: Science and Mind are mirror images of each other which use specific calculations over certain kinds of numbers. Mathematics offers an unbounded variety of calculating possibilities. The selection that we describe below is based on what *works* for Physics today. The future will always add new items on the open-ended list...

On mathematical language
Even though the presentation makes use of mathematical symbols by necessity and refers to some theory, the focus is put on epistemology. The reader is encouraged to skip all technical obscurities to concentrate on the philosophical content of the message delivered by computation. The point of view about mathematics that is presented here is the traditional image of an organised *system of knowledge*, rather than the post-modern view of a haphazard heap of concepts and statements. Mathematics is viewed as a vital part of the Philosophy of Mind. The Mind computes the imago mundi of each human being, the reservoir of information coming from inside and outside which make sense for itself. Computed results which cannot be interpreted at any given time are just ignored. And action is taken (at a conscious or unconscious level) on the basis of a *chosen* subset of information from the reservoir. There is always a *choice* even when the computing agent is not aware of his choice.

Numbers in Mind

Three types of numbers
The elementary building blocks of all computations taking place today inside scientific computers are of three types only, namely scalars, vectors (i.e. lists of scalars of length n) and matrices (i.e. arrays of scalars of size $n \times m$), where n, m are integers. Such a classification is inductive because vectors or square matrices can be treated in computation practice as scalars.

About scalars
In general, scalars belong to a field, an algebraic structure defined by the two

basic operations $+$ and $\times$, in which the inverse operations (subtraction and division) are also defined. The scalars most familiar to school children are the integers in $\mathbb{Z}$ $(0, \pm 1, \pm 2, \cdots)$, and the rationals in $\mathbb{Q}$ (fractions n/m with n and m integers); later on students learn about reals in $\mathbb{R}$ (represented on the number line). In addition to these three types of fields, computer science has popularised the use of the binary field $\mathbb{Z}_2 = \{0, 1\}$ to represent integers in base 2 (rather than 10). Nature (in human translation) uses two other basic fields of scalars for its processing of information, each of which introduces a novelty by means of geometry. Natural scalars can also be complex numbers or quaternions.

The complex numbers in $\mathbb{C}$ are scalars in the form of real vectors with $n = 2$ real dimensions and the multiplication is defined by means of addition and a new operation known as conjugation (see Appendix $\mathbf{A}_1$). Finally the quaternions in $\mathbb{H}$ are scalar being real vectors with $n = 4$ real dimensions recursively constructed from $\mathbb{C}$. The novelty is that multiplication is now *not* commutative in $\mathbb{H}$: $x \times y \neq y \times x$ (see Appendix $\mathbf{A}_2$).

About vectors

Vectors of dimension n over a scalar field K live in a *linear* vector space K^n where they can be added: $\alpha x + \beta y$ is well-defined for $x, y \in K^n$, $\alpha, \beta \in K$. But they cannot be multiplied unless $n = 2^k$, $k \geq 1$.

When $n = 2^k$ and a multiplication is defined in K^n the linear vector space becomes a multiplicative algebra, where multiplication may be neither commutative nor associative: $(x \times y) \times z \neq x \times (y \times z)$. Three infinite sequences of multiplicative algebras consisting of real vectors of dimensions $n = 2^k$, $k \geq 1$ play a fundamental role in our current understanding of the phenomenological scheme of things.

i) Two families of Clifford algebras $Cl_{\pm,k}$, $k \geq 1$, in which multiplication is associative (see $\mathbf{A}_3$).

ii) The third family consists of Dickson algebras A_k in which multiplication is recursively defined, and stops being associative for $k \geq 3$ (see $\mathbf{A}_4$).

About matrices

The geometric and dynamic concept of a linear map from a vector space of dimension n into another of dimension m can be quantified by a rectangular array of scalars of size $n \times m$ which is called a *matrix*. When of the same size, matrices can add, but the product of A $n \times m$ by B $p \times q$ is possible only if $m = p$ to yield AB $n \times q$.

Square matrices of size $n \times n$, or equivalently of order n, form an associative multiplicative algebra where $n \geq 2$ is an arbitrary integer. Algebras of matrices defined over $\mathbb{R}$ or $\mathbb{C}$ can naturally process information by *spectral decomposition*. Each matrix of order n produces 1 or 2 vectors:
(i) All n complex *eigenvalues* form a vector in $\mathbb{C}^n$ always present. The sum (resp. product) of all eigenvalues is known as the *trace* $\mathrm{tr}\, A$ (resp. *determinant* $\det A$) of the matrix A.
(ii) When the matrix is *not* diagonalisable, its Jordan structure produces a second vector in the form of a sequence of $n-1$ bits (or binary digits, see $\mathbf{A}_5$). This global information can be stored locally at the level of any eigenvalue λ, $1 \leq g \leq m$, $1 \leq l \leq m$, in the form of a $g \times l$ matrix Σ_λ with entries in $\mathbb{Z}_2$; Σ_λ reduces to a row (resp. column) if g (resp. l)$= 1$. As an example, let λ of multiplicity $m = 8$ have a Jordan form which consists of $g = 4$ Jordan blocks of respective sizes $l = r_1 = r_2 = 3$, $r_3 = r_4 = 1$, Then

$$\Sigma_\lambda = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = (\sigma_{ij}), \quad \text{where } r_i = \sum_{j=1}^{3} \sigma_{ij}, \ \sum_{i=1}^{4} r_i = m = 8.$$

A *rectangular* matrix ($n \neq m$) has no eigenvalues but it enjoys a Singular Value Decomposition (SVD) producing a vector with $\min(n,m)$ nonnegative components representing the *singular values* (see $\mathbf{A}_6$).

When $n = m$, any square matrix A enjoys the two decompositions (spectral *and* singular) simultaneously, producing 2 or 3 vectors in $\mathbb{C}^n$, $\mathbb{R}^{+n}$ and possibly $\mathbb{Z}_2^{n-1}$. In contrast with the Jordan form which is characteristic of the spectral *complexity* for A, the singular values carry only *metric* information about A. For example when A is invertible, its distance to singularity is measured by its least singular value which can be much smaller than the least modulus of its eigenvalues. Because square matrices can produce the information consisting of their eigen-/singular values, they can induce a *two-level* information processing when they are used as scalars, see below for more. This is true for the matrix of the multiplication map associated with any vector in a Dickson algebra.

Use of real vectors in dimension 2 or 4 is not the only possibility exhibited by computation to create scalars by induction. Numerical Linear Algebra indicates that square matrices can be thought of as macro-scalars enabling an

extremely efficient computation both theoretically and in practice [Chatelin (2012), Chapter 1][3].

Macro-scalars live in a noncommutative *matrix ring* whose computational potential exceeds by far that of the fields $\mathbb{R}, \mathbb{C}$ and $\mathbb{H}$. The dynamics of computation illustrates the *hierarchical* aspect of information theory which is found everywhere in experimental sciences.

Geometric Interpretation for Vectors that Multiply

Geometry within Clifford numbers

The explosive variety of the possible realisations of mathematical structures can be tamed by considering isomorphisms which are transformations preserving the essence of the various structures. By means of this classical reduction by isomorphism, E. Cartan discovered in 1908 a remarkable property for the two families $Cl_{\pm,k}$: they exhibit the same 8-periodicity property represented in Table 1 of $\mathbf{A}_3$. There are 3 common structures for $k = 0, 4$ and 8. In between, there are 12 distinct structures, yielding a total of 15 structures at play. For $8 \leq k \leq 16$ the same evolution takes place for each individual element of the real matrix $\mathbb{R}(16)$ of order 16 consisting of 256 real coefficients. And this continues forever.

By a close inspection of Table 1 in $\mathbf{A}_3$, we can discern four different dynamical patterns in action, each acting in two steps. For example the pattern $P_1 : \ S \to S^2 \to S(2)$ is visible for $Cl_{-,k}$ with $S = \mathbb{H}$, $k = 2,3,4$; it is visible for $Cl_{+,k}$ with $S = \mathbb{R}$, $k = 0,1,2$ The two patterns P_1, P_2 increase the algebraic complexity, whereas P_3 decreases steadily the complexity of the scalar fields for the matrix coefficients $(\mathbb{H} \to \mathbb{C} \to \mathbb{R})$, while the geometric dimension grows.

By imposing *associativity* to Clifford algebras, we obtain an increase in computational complexity which obeys a *finite* rule of evolution in 8 steps. Each of the *three* scalar fields $\mathbb{R}, \mathbb{C}$ and $\mathbb{H}$ plays a role: Cartan's unexpected result was instrumental in the belated acceptance of noncommutative quaternions by the physics community. Globally, the matrices may have the order $n = 2, 8, 16$ over $\mathbb{R}$, $n = 2, 4, 8$ over $\mathbb{C}$ and $n = 2, 4$ over $\mathbb{H}$. Table 1 reveals the ruling power of geometry to extract order from variety. This explains why

[3] It is not a new idea to replace a field by a ring as the algebraic structure for a set of scalars. But its most convincing application is ignored by algebra textbooks. The idea is (implicitly) at work in the fast matrix algorithms tailored for parallel architectures which underlie the boom of information technologies (from financial products to the web).

geometry has played an increasing role in Theoretical Physics during the past two centuries. Spinor groups are well chosen subgroups of units of Clifford algebras. Their (algebraic) action on $\mathbb{R}^k$ explains the rotations in k dimensions [Gallier 2011]. And, to the amazement of the physics community, the very same concepts serve to describe the interaction between matter and forces in the Standard Model of particle physics [Baez 2002, pp. 163-164].

Hypercomputation in Dickson algebras

Classical calculus relies heavily on derivatives to model and analyse the evolution of phenomena viewed as functions of certain real or complex variables. In a multiplicative algebra context, Leibniz's idea of a differential takes the form of a *derivation* (see $\mathbf{A}_7$). The limit of the algebraic reducibility by linear derivation in A_k is expressed by the nonlinear core $K(A_k)$ of dimension limited to 1 or 2 for $k \leq 3$, and equal to $(1/8)\dim A_k$ for $k \geq 3$. Table 2 in $\mathbf{A}_7$ indicates that the nature of the explicative variable for hypercomputation phenomena taking place in A_k can be *recursively* reduced to be real or complex with algebraic depth $= 1$ or 2. The two commutative fields $\mathbb{R}$ and $\mathbb{C}$ form, together with ∞, the set called *Reason* in hypercomputation [Chatelin 2012, Chapter 3].

Because of lack of freedom in the definition of $\times$ (see $\mathbf{A}_4$) nonassociative Dickson algebras exhibit much less geometric regularity than their free associative Clifford counterparts. But it goes against conventional wisdom that they should offer a much richer variety of computational opportunities [Chaitin-Chatelin 2007]. This counter-intuitive property can be revealed by the singular values of the left (say) multiplication map L_x defined by a vector x, that is $L_x : \ y \mapsto x \times y$, or equivalently the eigenvalues of $L_x^T L_x$.

For $k \leq 3$, $L_x^T L_x = N(x) I_{2^k}$: the unique singular value of L_x is $\|x\|$. The property does not hold in general for $k \geq 4$, however the weaker equality $\mathrm{tr} L_x^T L_x = 2^k N(x)$ remains true.

Vectors with a zero-component on the real *and* complex units 1 and $\tilde{1}$ are called doubly pure. Such vectors t yield $L_t^T L_t = -L_t^2$ with a remarkable eigenstructure: all eigenvalues are nonnegative with multiplicity $4p$, $p \geq 1$. The eigenvalue $N(t)$ is always present with $p = 2$ or 4 when $k = 4$.

In general $x = \alpha + \beta\tilde{1} + t$, with $h = \alpha + \beta\tilde{1}$ (head) and t doubly pure (tail); thus $L_x^T L_x = N(h) I_{2^k} - L_t^2$. The author discovered in 2005 that when the chosen singular value for L_t differs from 0 or $\|t\|$, there are three possible answers for the corresponding singular value of L_x which depend on the computational route [Chaitin-Chatelin 2007]: the answer is not anymore unique for $k \geq 3$, but becomes *threefold*.

Let us consider the above addition for x, which can be equally written as

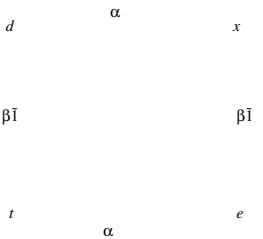
$$x = \alpha + (\beta\tilde{1} + t) = \alpha + d \qquad (*)$$

$$x = \beta\tilde{1} + (\alpha + t) = \beta\tilde{1} + e. \qquad (**)$$

Unexpectedly, the results for the local SVD on L_x computed by (*) and by (**) *differ* when $\alpha\beta \neq 0$. There are *three* possible routes to go, for $\alpha\beta \neq 0$, from t to x in the plane spanned by 1 and $\tilde{1}$:

1. directly by: $x = h + t = (\alpha + \beta\tilde{1}) + t$,
2. through d by (*),
3. through e by (**).

The three computational routes are sketched below. They coalesce into *one* route only when $\alpha\beta = 0$.



According to elementary matrix theory, any eigenvalue λ for $-L_t^2$ should produce the eigenvalue $N(h) + \lambda$ for $L_x^T L_x$. This is always satisfied for $\lambda = 0$ or $N(t)$, but this exact value represents only *half* of the possible values computed locally when $0 < \lambda \neq N(t)$.

In other words the addition $x = \alpha + \beta\tilde{1} + t$ appears nonassociative from the point of view of SVD computation.

Is the local SVD derivation absurd?

From the point of view of classical logic, half of the local SVD results are plainly **wrong** when they do not agree with the exact value. Moreover, when $\beta^2 = \alpha^2 + \lambda$, they produce pseudo-zerodivisors which contradict the theoretical property that true zerodivisors are necessarily doubly pure. At face value, local SVD seems **absurd**, and it should be rejected by any sane mathematician. Or should it not?...

Should we not think twice before taking action? In the 16th century, $\sqrt{-1}$ was a complete mystery, which appeared totally absurd at first sight.

It took three centuries of painful reflections by some of the greatest minds like Euler and Cauchy to master its meaning as the "imaginary" unit i. Once tamed, $i = \sqrt{-1}$ found its way in almost all engineering calculations of the 19th century which dealt with wave propagation (light, sound, electricity,magnetism,...). Warned by history, we should be extremely cautious... We should not jump hastily to the "obvious" negative conclusion. We should examine the possibility that local SVD derivation *serves a purpose* from a computational point of view, and that it delivers useful *information*. We should remember that during the 18th century, the resolution of the paradox created by $i = \sqrt{-1}$ was realised by *increasing* the dimension n of the number space from $n = 1$ (real line) to $n = 2$ (complex plane). Similarly, the paradox of the pseudo-zerodivisors in A_k can be partially explained in $A_{k+1} = A_k \oplus A_k \times \tilde{1}$. The dimension $n = 2^k$ should be doubled to become $2 \times 2^k = 2^{k+1}$. This complexification of the computing context has the remarkable property to give a theoretically valid meaning for the nonconforming SVD results which defy theory in the context of A_k alone. It also gives to the computing agent the *freedom to choose* between several options.

This **computational paradox** reveals a clash between the associative matrix algebra frame for computation and the truly nonassociative structure A_k, $k \geq 3$. This clash is extremely beneficial for computation. This is one of the mechanisms by which novelty can show up and exhibit the *creative* power of mathematical computation. The insistence on sticking to the classical logic of theory would deprive its proponent of some extremely useful tools to anticipate by accessing to a larger reality. The fixed logic option bans *free will* from computation and replaces *freedom of choice* by necessity for the computing agent.

We add that the three possible sets of singular values for L_x derived from those of L_t modify the local 3D-geometry at x (Chapters 5 and 6 in [Chatelin 2012]). The flexibility of the perceived geometry reflects the dynamical logical evolution of computation. This essential property of Life's computation has not yet received the serious attention it deserves.

The Nonassociative Octonions

When quaternions were discovered by Hamilton in 1848, their noncommutativity dissuaded most mathematicians and physicists to venture any further. Despite the strong support of Maxwell, 4D-quaternions were quickly expelled

from physics and replaced by the more familiar 3D-vectors. This exile would last for more than seven decades!

If noncommutativity was deterrent, nonassociativity was literally *unthinkable* then! It was not possible to make sense of the discovery by Graves[4], a classmate and friend of Hamilton, of the *octonions* – only three months after the momentous invention of quaternions. Minds were not ready yet for such an unruly multiplication of vectors in $\mathbb{R}^8$, refusing to be either commutative or associative [Hamilton (1848)].

Applications of quaternions to physics were challenging possibilities by the end of the 19th century (at work in electromagnetism and implicitly in Special Relativity). But nobody could foresee any conceivable role for the octonions. They would remain a seemingly useless concept until the dawn of the 21th century[5]. Some 150 years would be necessary for the scientific community to feel comfortable enough with the concept of nonassociative multiplication [Baez 2002, Baez and Huerta 2011].

The eventual come back of octonions into physics was nevertheless predictable. The algebra $A_3 = \mathbb{G}$ has *three* basic generators $g_1 = e_1$, $g_2 = e_2$, $g_3 = e_4$ which can be thought of as defining the familiar 3D-space in which we live (see $\mathbf{A}_4$, [Chatelin (2012)]). This multiplicative interpretation is much less obvious than the linear one. The role of $\mathbb{H}$ is visible in geometry (Table 1), whereas the role of $\mathbb{G}$ is more subtle to detect because of its inherent nonassociativity. This, of course, makes $\mathbb{G}$ a notion which reaches deeper into the "soul" of computation than $\mathbb{H}$. And the latest developments of String Theory in Theoretical Physics testify to this greater depth [Baez 2002].

The reader should keep in mind that, according to the point of view developed in this Chapter, the space we live in cannot have more than *three* physical dimensions. In 1907 Minkowski added one extra-dimension to the 3D-space to explain Special Relativity. Today up to 8 extra-dimensions are called for into theoretical accounts of particle physics. These new dimensions are ideas living in the minds of theoretical physicists: they serve to build coherent theories which are meant to open a rational window on the larger reality. Because Reality is inexhaustible, no theory will ever be final. Through mathematical

[4]The letter $\mathbb{H}$ classically used to represent A_2 is a tribute to Hamilton. Similarly to represent A_3, the letter $\mathbb{G}$ honours J. T. Graves, the true inventor of the octonions which are often wrongly attributed to Cayley in the scientific literature.

[5]A clear illustration of the original inability of the scientific mind to **think** of a nonassociative multiplication is provided by Bourbaki. This distinguished group of algebraists never felt the need to give a *specific* name to a nonassociative structure which would be a field if it were associative.

computation in *mind* and over *time* Life will always add new mental dimensions to the *evolving* scientific imago mundi.

Noncommutativity in Matrix Algebra

The noncommutativity of the product of two square matrices A and B, such that $AB \neq BA$ is usually not a severe computational hindrance. Its most obvious consequences are mainly reported as exercises in textbooks. This may be the reason why the identity (2) in $\mathbf{A}_8$ which links the resolvents for $E = AB$ and $F = BA$ has gone *unnoticed*. But a spectral analysis of the connection (2) near the eigenvalue $\lambda \neq 0$ by means of (1) in $\mathbf{A}_5$ leads to an unbounded sequence of conditions, of which $l+1$ have tangible computational consequences which take the general form (3) given in $\mathbf{A}_8$.

When E and F are not only similar but invertible $(\det AB \neq 0)$, $P_F = \frac{1}{\lambda} BP_E A$ is true, leading to a contradiction with (3b) when $l > 1$ ($D_F \neq 0$). When λ is not defective the two conditions (3a) and (3b) imply that $AB = BA = \lambda I$, $\lambda \neq 0$: A and B *commute* and λ is the *unique* eigenvalue of their product; λ is also the product of any eigenvalue μ for A by its corresponding eigenvalue $\frac{\lambda}{\mu}$ for B. This spectacular result puts **commutativity** under the spotlight. It shows that spectral analysis is an inner mechanism by which a special type of commutativity can *emerge* in a noncommutative matrix algebra. This confirms the well-known fact that a notion is best understood by looking at its absence.

In the general case, E and F are not invertible, and not necessarily similar. Then the infinite sequence of conditions expressed by (3) and (4) deserves special attention. It reveals the full spectral connection between AB and BA in the neighbourhood of $\lambda \neq 0$. When $\lambda = 0$, theory tells us that the indices l_E and l_F for 0 satisfy $l_E \in \{1,2\}$ for $l_F = 1$ and $l_E \in \{l_F - 1, l_F, l_F + 1\}$ for $l_F \geq 2$, cf. $\mathbf{A}_7$. Only additional information about E and F (such as numerical experimentation) can indicate which possibility is actually the case. As far as we know, the relations (3) and (4) have escaped the attention of analysts. The global connection (2) which exists between the noncommuting square matrices A and B $(AB \neq BA)$ evokes irresistibly the nonlocality property displayed by particles which has been the source of so many puzzling questions in Physics. It also provides a rational basis for the many *holistic* aspects of Life which are more and more frequently observed in experimental biology and ecology.

The above discussion indicates that a local spectral analysis of resolvents

for AB and BA may produce computational difficulties running into contradiction with the global connection (2) between these resolvents, when $AB \neq BA$. The fact that lack of commutativity may induce paradoxical conclusions by means of a spectral analysis appears to be *new* knowledge with serious epistemological consequences. To provide a better understanding of the potential significance, we recall that, when $AB = BA$ the product of the matrix exponentials e^A and e^B satisfies $e^A e^B = e^B e^A = e^{A+B}$. In other words, the commutativity of matrices *extends* to their exponentials. The converse is true if the matrices A and B are defined over the field of *algebraic* numbers (i.e. the algebraic closure of the integers). Such numbers are all roots of polynomials with *rational* coefficients.

It is quite remarkable that the above theory extends to *rectangular* matrices A and B^T of size $n \times p$ such that $E = AB$ of order n and $F = BA$ of order p are well-defined and *spectrally connected*, cf. $\mathbf{A}_8$. The two matrices AB and BA of different orders are actually strongly connected by a spectral analysis near any of their eigenvalues. The eigenvalue 0 plays a specific role in the spectral connection.

Local Vs. Global Information

The theoretical prediction of nonlocality in Quantum Mechanics, experimentally confirmed much later (Aspect 1982), has stirred many philosophical discussions about the dual nature of the observable outer reality.

Not too surprisingly most quantum physicists believe in the primacy of the quantum world to "explain" the one we observe around us. And this opinion is repeatedly confirmed by the variety of electronic devices which today fill our environment. The technological feats should not overshadow the other side of the coin: Quantum Mechanics is part of but not the whole story, as most of us are ready to acknowledge.

The clash between the local and global aspects of computation is a recurring phenomenon in Mathematics. We have seen that it can impact the local 3D-geometry at any vector x in nonassociative Dickson algebras A_k, $k \geq 3$ (based on the SVD for L_x). And in any associative matrix algebra, this clash can be revealed by *two* noncommuting invertible matrices. If $AB \neq BA$, a local spectral analysis of the two defective matrices AB and BA challenges their global connection. The difference in local/global information can also be revealed on a *single* square matrix A $n \times n$ by the Jordan structure of its spectral decomposition. Globally the structure can be represented by a *vector* in $\mathbb{Z}_2^{n-1}$.

Locally at an eigenvalue λ, $1 \leq g, l \leq m$, the structure can be represented by the *matrix* Σ_λ $g \times l$ defined over $\mathbb{Z}_2$, which produces information by means of its SVD in general, and also by its eigenvalues when $l = g$.

So far in this Chapter the focus of attention has been put on *multiplication*. We have described the most striking consequences for epistemology of the lack of commutativity and/or associativity for multiplication. The story continues with the other basic operation for computation which is *addition*.

On Relativistic Additions

We have seen above that the SVD computation for the multiplication map L_x defined by a dicksonian vector $x = \alpha + \beta\tilde{1} + t$ challenges the associativity of the addition of vectors in A_k, $k \geq 3$. But Nature provides us with the more radical modification of addition illustrated by Einstein's relativistic law of addition for velocities near the speed of light (1905).

It was discovered in [Ungar (1988)] that Einstein's peculiar addition of velocities could be interpreted more generally as a nonstandard addition of vectors (denoted $\oplus$) whose departure from commutativity and associativity is controlled by a rotation. The additive formula is *relativistic* in the following sense. Let be given a reference value $\lambda > 0$ which defines the ball $B_\lambda = \{x \in \mathbb{R}^n, \|x\| < \lambda\}$. Then for two given vectors x, y in $\mathbb{R}^n$, the output of $x \oplus y$ consists either of a unique *real* vector when $\|x\| < \lambda$ or of a pair of *complex* conjugate vectors when $\|x\| > \lambda$. The sum $x \oplus y$ can only represent a *physical* velocity if it is a *real* vector. In Special Relativity, the reference value is taken to be $\lambda = c$ the speed of light, and $\oplus$ applies to the velocity of fast (high energy) particles. This motivates the physical principle that no dynamical information can travel faster than light.

The relativistic addition modifies the global euclidean geometry of $\mathbb{R}^n$ inside the ball B_λ. Einstein's addition underlies the Beltrami ball-model (1868) of hyperbolic geometry of $\mathbb{R}^n$. It is remarkable that the Beltrami model of 1868 preceded by almost four decades its physical illustration by Einstein!

Hyperbolic translation is the relativistic addition of Poincaré (related to conformality and incompressible fluid flow); it underlies the Poincaré ball-model (1880) for the same hyperbolic space, see $\mathbf{A}_9$. The two models are the two geometric expressions of the relativistic additions independently studied by Einstein and Poincaré which have the same technical origin in the group of automorphisms of the complex unit disc in $\mathbb{C}^n$ [Ungar 1997, Chatelin 2011d].

This common computational origin in *complex* analysis exposes the futility of the academic opposition between Einstein and Poincaré which can often be found in historical accounts of Special Relativity.

The mathematically inclined reader in want for more may look at the reports [Chatelin 2011b,c,d] which are strewn with epistemological remarks.

Hyperbolic geometry in Nature
A number of natural shapes exhibit, at least locally, a hyperbolic character in their geometry. The most famous example is a horse saddle or a mountain pass. Among other natural hyperbolic surfaces, one can cite lettuce leaves, coral reef or some species of marine flatworms with hyperbolic ruffles. If one moves away from a point in hyperbolic plane, the space around the point expands exponentially. The idea was implemented in crochet in 1997 by D. Taimina by ceaselessly increasing the number of stitches in each row of her crochet model [Henderson and Taimina 2001]. Experiments have shown that the visual information seen through the eyes and processed by our brain is better explained by hyperbolic geometry (Luneburg 1950). This explains the current popularity of hyperbolic browsers among information professionals. And hyperbolicity underlies *fractality*, a distinctive property of many natural objects, from clouds and trees to coastal lines.

Mechanical Computation

Our description of mathematical computation would not be complete if we omitted the very specialised field known in Theoretical Computer Science as "Computability Theory" (also called "Recursive Function Theory" by logicians). Following the programmes of Hilbert and Turing in the 1930s, the theory develops a *finitist* version of computation adapted for machines where infinite processes have no place: the infinite binary sequence $0,111\cdots=\sum_{1}^{\infty}2^{-k}$ is Turing-computable but cannot be identified with the number 1. This modern ban of *actual* infinity by Theoretical Computer Science is the radicalisation of the caution and philosophical reservation it provoked in Ancient Greece. But Archimedes knew very well how to sum certain geometric series by exhaustion!

Such a radical viewpoint is not espoused by all logicians In Section 3 of [Gödel (1972), p. 306] entitled "A philosophical error in Turing's work", Gödel explains why the mechanical thesis of Turing *cannot* fully apply to

human thinking in the following terms: "What Turing disregards completely is the fact that *mind, in its use, is not static, but constantly developing*, i.e., that we understand abstract terms more and more precisely as we go on using them, and that more and more abstract terms enter the sphere of our understanding." (italics in the original). Gödel goes on by stressing the necessary role of ∞, showing a clear understanding that $\{\mathbb{N}, \infty\} \subset$ Reason.

It is perplexing to witness that Gödel's warning against the risks of a naive extrapolation from machine to mind is forgotten by so many logicians and neuro- as well as computer scientists in the 21st century.

In order to model the perceived evolution of the mind toward abstraction, modern logicians have followed Euclid's footsteps in starring the concept of *axiom*: they propose to add new axioms to the agreed-upon basis of ZFC (Zermelo-Fraenkel-axiom of choice) for logic. ZFC can be viewed as a formal logic version of $\{\mathbb{Z}, \infty\}$. And stronger axioms of infinity in cantorian set theory represent an attempt to realise the Rational Core $= \{\mathbb{R}, \infty\}$ (which misses $\mathbb{C}$ from a purely algebraic viewpoint). The notion of an axiom has some merit for the one dimensional mental activity known as thought, expressed in a *sequential* language for communication. However it plays no fundamental role in intuition, a definitely two dimensional mental activity which allows logic to evolve in the mind. Actually any purely axiomatic approach forbids logical evolution. Therefore it runs the risk to kill mathematical invention; it may leave room for sterile abstraction rather than true discovery.

The logical claim that Turing machines are universal computers is valid only if one limits severely the definition of a computer to be a machine working with a logic that is *a priori fixed* and with numbers of algebraic depth 1 only [Chatelin 2012, Chapters 3 and 8].

It is clear that the human mind is a computer whose computing potential (but not speed!) exceeds by far that of a Turing machine. The human mind can access (at least implicitly) to infinity and to the uncountable, and its rational logic is stretchable.

Conclusion

Our journey into the land of mathematical computation has been *selective* in an effort to stay close to what could be a computational toolkit for a physicist of today. One essential lesson that mathematical computation is teaching us is that multiplication is the *lead actor* within the general information processing over numbers which can be scalars, vectors of dimension $2^k \geq 2$ or

square matrices of arbitrary order $n \geq 2$. Meaning is woven through mathematical computation by an interaction of the three classical domains of algebra, analysis and geometry. The computational process may force the logical mind of scientists to *expand* and accept new possibilities as valid results for life even if they defy the currently agreed upon logic. History tells us that, thanks to the evolution of logic, our predecessors could accept – after centuries of debate and inner anguish – such numbers as $\sqrt{2}$, 0 and $\sqrt{-1}$ which were highly problematic in their days [Chatelin 2012, Chapter 1]. Thanks to computers, numerical analysts have realised in the 1980s that, in computing practice, scalars need not enjoy a field structure. They can belong to a matrix ring and participate in a hierarchical information processing by computation. One usually does not think of matrices as *numbers*. Therefore it is easy to overlook the irreplaceable role played by such "numbers" as elementary *local sources* of information during any *global* processing of information.

It is extremely significant that paradoxical consequences of the lack of associativity or commutativity for multiplication are revealed by a spectral/singular analysis of matrices in $\mathbb{C}$ or $\mathbb{R}^+$ which in turn influences the geometric picture, either globally or locally. The difference between local and global perspectives can even be felt, beyond geometry, in the spectral/singular information provided by each Jordan box associated with an arbitrarily chosen distinct eigenvalue for a square matrix.

Computation has been the driving force behind the superb development of mathematics during some four millennia. But it has gradually been outmoded during the 20th century by the structuralist and formalist approaches which still rule mathematical fashion. May be time has come to reconsider the current preeminence of form and structure over computation. Without denying the importance of the former two notions which are static, it does seem that there is even more to be learnt from the dynamical mechanism of computation. It can tell us a lot about the ways by which the human mind processes numbers to construct the personal imago mundi that each of us uses to navigate in the reality we individually perceive and collectively build on earth at the same time.

The modern view of mathematics is that of an intimidating majestic building resting on solid logical foundations. To some it even looks like a giant tautology!

This view is directly inherited from the logical revolution of the 20th century which dogmatically promulgated the preeminence of logic over compu-

tation in the 1930s. This formalist point of view has some merit because it strengthens the logical coherence of the whole mathematical enterprise. But it relegates meaning in the background. The logical validity of a statement is a necessity which should not outweigh its epistemological value. As a predictable outcome of the primacy of logic over meaning, mathematical computation – so brilliantly served by Euler, Cauchy and Jordan – became disreputable; it is considered today by many mathematicians as a mere ancillary necessity better to be delegated to machines by means of computer algebra software...

However logic is the servant of *creative* mathematics. An excess of logic freezes the inventive spirit of mathematics: logic cannot be the ruling queen in the quest for knowledge. The evolution of logic goes hand in hand with the evolution, under computational pressure, of mathematics as a philosophical system of knowledge. For example, the dynamics of computation reveals the inductive necessity to welcome square matrices into the scalar toolbox of the scientific mind. This necessary evolution is not yet generally acknowledged. The illuminating description given by Gödel of the *evolutive* nature of logical understanding remains today a brilliant exception. Mathematics is an on-going process taking place in the minds of mathematicians, with its source in questions raised either from outside by the physical outer reality, or from inside by the inner reality of the mind and of life as we perceive it. The scientific divide between in and out is only a methodological artifact. When it was first conceived of by Galileo and Descartes, it was a bold step forward on the part of these deep thinkers. But in conjunction with material progress, this viewpoint has also brought to the western man a counter-productive alienation from a *working* understanding of life. More and more scientists become now painfully aware of this disappointing outcome. The global crisis that the world is experiencing now points sorely to the limits of this dichotomy. The harsh 20th century has shattered all naive materialist expectations: no amount of electronic gadgets will ever bring the vital satisfaction and joy that anyone longs for in the heart of his/her hearts.

If the goal that humanity sets for itself is collective peace, individual serenity and harmonious cooperation with all natural lives on the planet then it is time for the pendulum to swing back and for science to embrace a more balanced view of the mind-body split. The subjective aspect of life cries out for a scientific attention in which mind and matter cooperate to evolve the perceived reality. It is time for a "re-enchantment of the world". Nondogmatic

mathematical computation could be a big help in this endeavour.

Appendix

A$_1$ *The field* $\mathbb{C}$
In $\mathbb{C}$, $z = a + ib = (a,b) \in \mathbb{R}^2$, $i = \sqrt{-1}$, Arg $i = \pi/2$, $\bar{z} = (a,-b) = a - ib$, $zz' = (a,b)\cdot(a',b') = (aa' - bb', ab' + a'b)$,

$i = (0,1)$ is the complex unit in $\mathbb{C}$. Complex numbers fill a plane and lose the natural order enjoyed by real numbers on a line.

A$_2$ *The skew field* $\mathbb{H}$

$$\begin{aligned} \text{In } \mathbb{H},\ q &= (z_1, z_2) \in \mathbb{C}^2 \\ &= a_1 + \boldsymbol{i}b_1 + (a_2 + \boldsymbol{i}b_2)\times\boldsymbol{j} = a_1 + b_1\boldsymbol{i} + a_2\boldsymbol{j} + b_2\boldsymbol{k},\ \boldsymbol{i}\times\boldsymbol{j} = \boldsymbol{k}. \end{aligned}$$

$q \times q' = (z_1, z_2)\times(z_1', z_2') = (z_1 z_1' - \overline{z'}_2 z_2, z_2' z_1 + z_2 \overline{z'}_1)$.

$\boldsymbol{i} = (01,00) = (i,0)$, $\boldsymbol{i}\times\boldsymbol{j} = -\boldsymbol{j}\times\boldsymbol{i}$,

$\boldsymbol{j} = (00,10) = (0,1)$ is the complex unit in $\mathbb{H}$. Quaternions are best suited to represent spatial rotations. They are irreplaceable tools in certain engineering domains, from orbital mechanics for GPS to 3D-graphics for the film industry.

A$_3$ *Clifford algebras*
Clifford algebras are associative algebras freely generated by a multiplication obeying the following rule on k given vectors $\{i_l\}$, $l = 1$ to k. They should satisfy $i_l \times i_m = -i_m \times i_l$ for $l \neq m$ and $i_l \times i_l = \varepsilon \|i_l\|^2$, where $\|\cdot\|$ denotes the euclidean norm and $\varepsilon = \pm 1$. The resulting algebras denoted $Cl_{\pm,k}$ consist of Clifford numbers in a vector space of dimension 2^k defined over $\mathbb{R}$. It is important to observe that the multiplicative operation is largely *arbitrary*.

Let S represent an algebraic structure, we shall denote by $S^2 = S \oplus S$ the 2D-space of pairs (s,s'); if S is a scalar field, then $S(n)$ denotes the square matrix of order n with coefficients in S.

The period 8 table for $Cl_{\pm,k}$ discovered by E. Cartan (1908) is given by

Table 1. [Baez 2002, Gallier 2011]

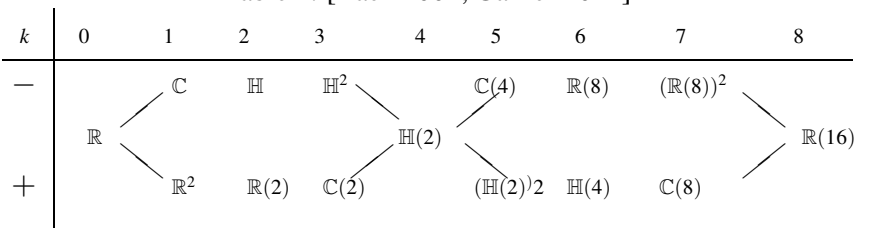
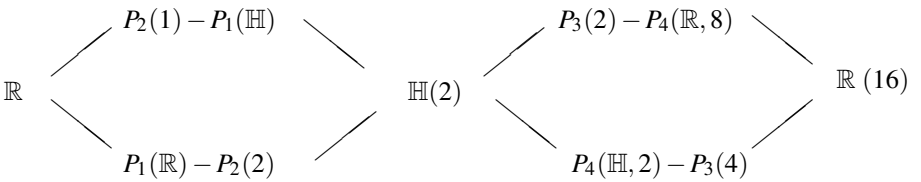


Table 1 displays four dynamical patterns, each consisting of two steps:

- $P_1(S):\ S \to S^2 \to S(2)$
- $P_2(n):\ \mathbb{R}(n) \to \mathbb{C}(n) \to \mathbb{H}(n)$
- $P_3(m):\ \mathbb{H}(m) \to \mathbb{C}(2m) \to \mathbb{R}(4m)$
- $P_4(K,n):\ S = K(n) \to S^2 \to K(2n)$

The patterns P_1 and P_4 share the same first step $S \to S^2$. The result of the second step in P_4 is $K(2n) \neq (K(n))(2) = S(2)$;it can be viewed as a *deconstruction* of that in P_1, a sudden *loss* of structure. This loss may appear as a source of randomness in computation.

Table 1 can be rewritten schematically as follows:



We observe that $\mathbb{C}$ does not appear in the realisation of P_1 and P_4.

A$_4$ *Dickson algebras*

The Dickson algebras A_k over $\mathbb{R}$ are defined recursively by the doubling process of Dickson: $A_{k-1} \to A_k = A_{k-1} \times 1_k \oplus A_{k-1} \times \tilde{1}_k$ where $1_k = (1_{k-1}, 0)$ is the real unit and $\tilde{1}_k = (0, 1_{k-1})$ is the complex unit, $\overline{(x,y)} = (\bar{x}, -y)$, and multiplication is defined by: $(x,y) \times (x',y') = (x \times x' - \bar{y}' \times y, y' \times x + y \times \bar{x}')$.

The complex algebra A_k, $k \geq 1$ is defined in a twofold manner: *linearly* as a vector space by 2^k basis vectors $\{e_i\}$, $i = 0$ to $2^k - 1$, or *multiplicatively* as an algebra by k basic generators $\{g_j = e_{2^j}\}$, $j = 1$ to k.

These algebras are quadratic: $x \times x = -x\bar{x} + (x+\bar{x})x$ where $N(x) = x\bar{x} = \|x\|^2 = \sum_{i=0}^{2^k-1} x_i^2$ and $\{x_i\}$ denote the components of x. Observe that $N(x)$, the *square* of the euclidean norm $\|x\|$, is a most natural way to measure the vector x. Let x and y be orthogonal and $z = x + y$, then $N(z) = N(x) + N(y)$ is the familiar theorem of Pythagoras.

The rule $N(x)N(y) = N(x \times y)$ is always valid for $k \leq 3$, but not in general for $k \geq 4$.

Consequently, there may exist nonzero vectors x and y in A_k, $k \geq 4$, such that $x \times y = 0$: they are zerodivisors. When x is a zerodivisor, it enjoys an infinity of inverses $x^{-1} \in \frac{\bar{x}}{N(x)} + \{y; x \times y = 0\}$.

The first four Dickson algebras which permit *division* deserve special attention. The first three $\mathbb{R}, \mathbb{C}$ and $\mathbb{H}$ are associative and enjoy a *field* structure. The fourth one $\mathbb{G} = A_3$ is nonassociative but alternative; its structure is sometimes called a *loop* in physics and – very inappropriately – a nonassociative *ring* in mathematics! Actually $\mathbb{G}$ enjoys a noncommutative and nonassociative "field" structure. The four division algebras share the arithmetic property of admitting rings of *integers* with respective dimension $1, 2, 4$ and 8, in which the factorisation into *primes* is essentially unique (up to units, i.e. integers with unit norm). There are respectively $2, 4, 24$ and 240 units in the integer rings, whose geometric properties play a fundamental role in Physics [Baez 2002], and in Number Theory [Chatelin 2012, Chapter 9].

A$_5$ *The Jordan form of a matrix of order n over* $\mathbb{C}$

For any square A, there exists an invertible X such that $A = XJX^{-1}$ where the Jordan form J is either (i) diagonal or (ii) bidiagonal (Jordan (1870)). The diagonal elements of J are the *eigenvalues* of A which form the *spectrum* $\sigma(A)$.

The latter case (ii) arises as soon as one distinct eigenvalue λ for A has g independent eigenvectors with $g < m$, the algebraic multiplicity of λ as a root in $\mathbb{C}$ of the characteristic polynomial $\det(zI - A) = 0$; λ is said to be *defective*. Each of g independent eigenvectors x_i, $i = 1$ to g can start a Jordan chain of length r_i, $1 \leq r_i \leq m$ consisting of vectors related by

$$x_{i0} = x_i,\ Ax_{ij} - \lambda x_{ij} = x_{i(j-1)},\ j = 1 \text{ to } r_i - 1, \text{ when } r_i \geq 2.$$

Moreover $\sum_{i=1}^{g} r_i = m$. Then the m vectors $\{x_{ij}, j = 0 \text{ to } r_i - 1, i = 1 \text{ to } g\}$ form a basis for the invariant subspace of dimension m associated with λ. Any choice in ordering the Jordan blocks for λ specifies the corresponding Jordan

structure, that is the binary sequence of length $m-1$ which displays $g-1$ zeros separating subsequences of $r_i - 1$ ones when $r_i \geq 2$.

Two matrices A and $B = XAX^{-1}$ are *similar*: they represent the same endomorphism on the vector space in *different* bases. Similar matrices belong to the same equivalence class characterised by their common Jordan form. It follows that isospectral matrices which are *not* diagonalisable need not be similar.

Let λ be a defective eigenvalue of A, $g < m$ with $1 < l = \max_i r_i \leq m$, which is isolated from the rest of the spectrum by a closed Jordan curve Γ. The resolvent matrix $R(z) = (A - zI)^{-1}$ is defined for $z \in \Gamma$ and $P = \frac{-1}{2i\pi} \int_\Gamma (A - zI)^{-1} dz$ satisfies $P^2 = P$: it is the *spectral* projection on the *invariant* subspace for λ. Set $D = (A - \lambda I)P$, then $D^k \neq 0$ for $1 \leq k \leq l-1$ and $D^l = 0$. Finally, define $S = \lim_{z \to \lambda} R(z)(I - P)$.

Then for z in the neighbourhood of λ the series

$$R(z) = -\frac{P}{z-\lambda} - \sum_{k=1}^{l-1} \frac{D^k}{(z-\lambda)^{k+1}} + \sum_{k=0}^{\infty} (z-\lambda)^k S^{k+1} \tag{1}$$

exists for $z \neq \lambda$ and defines a meromorphic function of z with a pole of order l at λ: $1 \leq l \leq m$ is the *index* of λ. The factor $z - \lambda$ appears with the exponent m (resp. l) exactly in the characteristic (resp. minimal) polynomial for A. When $1 < l < m$, the series (1) for $R(z)$ does not reflect the full algebraic complexity of λ. If $A = \lambda I$, then $P = I$, $D = S = 0$ and $R(z) = \frac{1}{\lambda - z} I$, $z \neq \lambda$. See [Chatelin 1993] for more.

$\mathbf{A}_6$ *The SVD of A of size $n \times m$ over $\mathbb{R}^+$*

Let A be of size $n \times m$, $n \geq m$ (say). When A is real (resp. complex) we define the transpose matrix A^T ($a_{ij} \to a_{ji}$) (resp. the conjugate transpose matrix A^H ($a_{ij} \to \overline{a_{ji}}$)) of size $m \times n$. Then the products $C = AA^T$ (resp. AA^H) and $D = A^T A$ (resp. $A^H A$) are square matrices with respective order $n \geq m$. Both C and D are symmetric (resp. hermitian) with respectively n and m non-negative eigenvalues: their spectra can differ only by the multiplicity of 0 (which is necessarily an eigenvalue of C when $n > m$). We denote by σ_i^2, $i = 1$ to m, the eigenvalues of D. The square roots $\sigma_i \geq 0$ are the *m singular values* of A. The singular form of A (Jordan (1874)) is the $n \times m$ matrix Σ with all elements 0 except for σ_i on the "diagonal" $1 \leq i = j \leq m$. There exists two orthogonal (resp. unitary) matrices U and V of respective order n

and m such that the unitary equivalence relation $A = U\Sigma V$ holds. In the real case $AA^T = U\Sigma VV^T\Sigma^T U^T = U\Sigma\Sigma^T U^T$ and $A^TA = V^T\Sigma^T\Sigma V$: the orthogonal matrices U and V define orthonormal eigenvectors for C and D respectively, corresponding to the eigenvalues on the diagonal matrices $\Sigma\Sigma^T$ and $\Sigma^T\Sigma$.

$\mathbf{A}_7$ *Linear derivation in* A_k
A derivation is a linear map D such that

$$D(x \times y) = (Dx) \times y + x \times (Dy) \text{ for } x, y \in A_k,\ k \geq 0.$$

The set of all derivations in A_k form the Lie algebra $\mathcal{D}er\ (A_k)$. The nonlinear core $K(A_k)$ is the subalgebra $\cap\{\text{Ker } D; D \in \mathcal{D}er\ (A_k)\}$ that is the set of nonzero vectors without an explanation by linear derivation; $d_a(k) = \dim K(A_k)$ is the *algebraic depth* of A_k, It is given (Eakin and Sathaye 1990) by the table:

Table 2.

k	0	1	2	3	$k \geq 4$
$d_a(k)$	1	2	1	1	2^{k-3}

$\mathbf{A}_8$ *Spectral analysis of AB and BA*
When A (*or* B) of order n is invertible, then AB and BA are similar: $A(BA)A^{-1} = AB$. When $\det A = \det B = 0$, AB and BA share the same characteristic polynomial: they are isospectral, hence they are surely similar when they are diagonalisable (for example if $B = A^T$ or A^H). Even if AB and BA are not similar, the following augmented matrices of order $2n$ are always similar:

$$\begin{pmatrix} I_n & A \\ 0 & I_n \end{pmatrix} \begin{pmatrix} 0_n & 0 \\ B & BA \end{pmatrix} \begin{pmatrix} I_n & -A \\ 0 & I_n \end{pmatrix} = \begin{pmatrix} AB & 0 \\ B & 0_n \end{pmatrix}.$$

The Jordan forms for AB and BA differ only at the eigenvalue 0 for which the sizes r_i of the Jordan blocks satisfy $|r_i(AB) - r_i(BA)| \leq 1$, $i = 1$ to $\max(g(AB), g(BA))$ while keeping $m = \sum_{i=1}^{g(AB)} r_i(AB) = \sum_{i=1}^{g(BA)} r_i(BA)$ invariant as the common multiplicity of 0 (Flanders 1951). Let 0 be a defective eigenvalue of AB characterised by the 3 integers g, l, m with $g, l \in [1, m]$ (see $\mathbf{A}_5$). Then, the structural matrix $g \times l$ for AB can be associated with finitely many possible structural matrices $g' \times l'$ for BA. The total number $N(g, l, m)$ of possibilities is

minimum at 1 when AB and BA are similar. When the similarity exists only at the augmented level $2n$, N can grow exponentially with m. Two special cases are important for natural evolution when $m \geq 2$. When g (resp. l) $= 1$, the structural matrix for $0 \in \sigma(AB)$ is a row $1 \times m$ (resp. column $m \times 1$) with *one* singular value $\sqrt{m}$. We suppose that the Jordan form for 0 *evolves* in $\sigma(AB)$. When $g = 1, l = m$ then necessarily $g' = 2, l' = m-1$. And when $g = m, l = 1$, then $l' = 2$ and g' can vary in $[p, m-1]$ for $2p = m$ or $m+1$. There are *two* distinct singular values for the new structural matrix which belong to $[0, \sqrt{m}]$; the end values $\{0, \sqrt{m}\}$ are obtained for $g' = p = \frac{m}{2}$ when m is even.

By exploiting the associativity of the matrix product $A(BA) = (AB)A$ we get, for any $z \notin \sigma(AB) = \sigma(BA)$ that $-z(BA - zI)^{-1} + B(AB - zI)^{-1}A = I$. Denote $E = AB$ and $F = BA$, the resolvents $R_E(z) = (E - zI)^{-1}$ and $R_F(z) = (F - zI)^{-1}$ are *connected* by the relation

$$BR_E(z)A - zR_F(z) = I = BR_E(z)A - (z - \lambda + \lambda)R_F(z) \tag{2}$$

at any $z \notin \sigma(E) = \sigma(F)$.

When $\det BA = (\det B)\det A \neq 0$, $BR_E(z)A = -\frac{1}{z}R_{F^{-1}}(\frac{1}{z})$ at $z \notin \sigma(F) \cup \{0\}$, hence (2) is equivalent to the tautological relation $\frac{1}{z}R_{F^{-1}}(\frac{1}{z}) + zR_F(z) = -I$. Because the function $z \mapsto \frac{1}{z}$ is defined on $\sigma(F)$, $P_F = P_{F^{-1}} = \frac{1}{\lambda}BP_EA$ for $\lambda \in \sigma(F)$.

In general, we can perform a *local* spectral analysis for z close to $\lambda \neq 0$ by replacing in (2) the resolvents by their series expansions (1). The matrix coefficients for $\frac{1}{(z-\lambda)^k}$, $k = 0$ to l, should vanish, yielding the $l+1$ relations (3):

- $k = 0$: $(I - P_F)(I + \lambda S_F) = BS_EA$, (3a)
- $k = 1$: $D_F + \lambda P_F = FP_E = BP_EA = 0$, (3b)
- $2 \leq k \leq l-1$: $FD_F^{k-1} = (F - \lambda I)^{k-1}BP_EA = BD_E^{k-1}A$,
- $k = l$: $\lambda D_F^{l-1} = (F - \lambda I)^{l-1}BP_EA$,

The matrix coefficient for $(z - \lambda)^k$, $k > 0$ should also vanish, yielding the infinite sequence of conditions (4):

$$BS_E^{k+1}A = (I + \lambda S_F)S_F^k, \quad k \geq 1.$$

Let A, B^T be $n \times p$ matrices, $n > p$ (say), with $E = AB$ of order n and $F = BA$ of order p. Then $\sigma(E) = \sigma(F) \cup \{0^{n-p}\}$. For $z \notin \sigma(E)$, $AR_F(z)B - zR_E(z) = I_n$.

A$_9$ *Two ball-models for the hyperbolic space*

After a long quest which lasted for several centuries, it was finally proved in the 19th century that the fifth axiom of Euclid is *optional* in geometry. Given a point A lying outside a line D which define together a 2D-plane, one may draw through A respectively *one* line parallel to D in euclidean geometry, *many* parallel lines in hyperbolic geometry and *none* in elliptic geometry. In the Belrami ball-model, hyperbolic geodetic lines are euclidean line segments, and in the Poincaré ball-model, hyperbolic circles are euclidean circles. The two dual models are, in a mathematical sense, "equivalent" and share the same computational source in complex analysis [Ungar 1997].

References

[1] R. Axelrod (1984). *The evolution of cooperation.* New York: Basic Books.

[2] J. Baez (2002). The octonions, *Bull. AMS*, *39*, 145–205.

[3] J. Baez and J. Huerta (2011). The strangest numbers in String Theory, *Scientific American, May 4*, 61–65.

[4] F. Chaitin-Chatelin (2007). Computing beyond classical logic: SVD computation in nonassociative Dickson algebras, in *Randomness and Complexity from Leibniz to Chaitin* (C. Calude ed.), pp. 13–23. Singapore: World Scientific.

[5] J. P. Changeux and A. Connes (1995). *Conversations on Mind, Matter, and Mathematics*. (M. DeBevoise ed. and transl.) Princeton, NJ: Princeton University Press.

[6] F. Chatelin (1993). *Eigenvalues of matrices*, Chichester: Wiley; updated version republ. 2012, SIAM Philadelphia.

[7] F. Chatelin (2011a). A computational journey into the mind. *NatComput Special Issue: UC 2010, (J. Timmis and K. Morita guest eds.)*, pp 1–13. doi:10.1007/s11047-011-9269-6, publi. *11*, 67–79 (2012).

[8] F. Chatelin (2011b). Polymorphic information processing in weaving computation: An approach through cloth geometry. Technical Report TR/PA/11/27, Cerfacs.

[9] F. Chatelin (2011c). On relating computation. Technical Report TR/PA/11/37, Cerfacs.

[10] F. Chatelin (2011d). About weaving computation in 2 and 3 real dimensions Technical Report TR/PA/11/119, Cerfacs.

[11] F. Chatelin (2012). *Qualitative Computing: a computational journey into nonlinearity*. Singapore: World Scientific.

[12] T. Dantzig (1954). *Number. The language of Science* (4th edition, revised and augmented). New York: The Free Press, Macmillan Publ.

[13] J. Gallier (2011). Clifford algebras, Clifford groups and a generalisation of the quaternions, e-manuscript, CIS Dept, U. Pennsylvania, Philadelphia.

[14] K. Gödel (1972). Some remarks on the undecidability results, In *Collected works II, Publications* 1938–1974 (pp. 305–306), Oxford: Oxford Univ. Press (1990).

[15] R. W. Hamilton (1848). Note respecting the researches of John T. Graves, Esq. *Trans, Royal Irish Acad. 21*, 338–341.

[16] D. Henderson and D. Taimina (2001). Crocheting the hyperbolic plane, *The Math. Intell. 23*, 17-28.

[17] E. Husserl (1970). *The crisis of European Sciences and Transcendental Philosophy*. (Carr D. ed. and transl.) Evanston: Northwestern University Press.

[18] E. Klein (2008). *Galilée et les Indiens*. Paris: Flammarion.

[19] H. Poincaré (1905). *La valeur de la science*, Paris: Flammarion; republ. 1970.

[20] A. Ungar (1988). Thomas rotation and the parametrization of the Lorentz transformation group, *Found. Phys. Lett.*, *1*, 57–89.

[21] A. Ungar (1997). Thomas precession: its underlying gyrogroup axioms and their use in hyperbolic geometry and relativistic physics. *Found. Phys.*, *27*, 881–951.

The Cerfacs reports are downloadable from http://www.cerfacs.fr/6-26641-Technical-Reports.php

INDEX

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y