5

HABERMAS AND THE KANT-HEGEL CONTRAST

Espen Hammer

Hegel's influence on the thinkers associated with the early Frankfurt School has been vast. The notion of immanent critique, which Max Horkheimer, Theodor W. Adorno and Herbert Marcuse appropriated and used in their reflections on contemporary society, was in large part generated from their reading of Hegel, and so was their commitment to offering a philosophical account of the nature and aspirations of modernity. Their insistence on interpreting cultural and social phenomena dialectically had Hegel's method of self-reflection in the *Phenomenology of Spirit* as its most immediate forerunner, and they have all in some defining sense subscribed to an essentially Hegelian notion of *critique*.[1] In particular, what has united their efforts and given Critical Theory a sense of identity has been the project of trying to unravel what they have seen as the contradictions of late modernity between its actual practices and its ideals, and then, on the basis of that unraveling, perform an immanent social critique aimed at indicating how those contradictions might be overcome. Hegel, they have argued, offers a view of social reality that is not fixated on social facts as such, as though these exist on a par with those of the natural world, but on the way in which such facts are constituted and generated in processes that themselves, though not exclusively, involve claims and commitments that the critical theorist can turn to in order to obtain normative resources for conducting critique.[2] Most of the early proponents of Critical Theory, including Horkheimer, Adorno and Marcuse, followed the early Marx, however, in criticizing Hegel for having developed what they saw as an excessive form of idealism.[3] Hegel, they argued, ultimately succumbed to the idealist temptation to provide final, and in the ultimate instance indefensible, groundings for his philosophy.

More recently, Jürgen Habermas has sought to reconceptualize the nature of Critical Theory by turning more directly to Kant. In a crucial move, which in the early 1970s led him to embark on the long path of developing a formal-pragmatic theory of rational communication, Habermas claimed that *critique*, the central term, obviously, of any *critical theory*, involves not only, as he had initially argued, self-reflection (or immanent critique) but also, in a more Kantian vein, rational reconstruction of necessary presuppositions of rational

action and judgment.[4] In contrast to the earlier representatives of Critical Theory, for whom reification and positivism (or, in Hegelian jargon, dogmatism) emerged as the fundamental issues to be confronted, Habermas's turn to Kant was in large part motivated by a growing desire to refute contextualism and relativism.

Echoing tendencies in thinkers such as Robert Brandom, Richard Rorty, Hilary Putnam, Stanley Cavell and Donald Davidson, what we might think of as the Hegelian side of Habermas is less interested in questions of representation (at least if conceived in strong realist terms, and with reference to an individual subject's experience independently of a community of speakers) than in how speakers, while belonging to a linguistic community, rationally go about understanding, communicating and justifying their claims to one another. Yet while reason's embodiment in language, practice and historical forms of life has become a dominant theme in contemporary post-analytical philosophy, the Kantian side of Habermas has called for the reconstruction of formal-pragmatic presuppositions of speech that can provide an account of how unconditional and context-transcending claims to universal validity are possible.[5] The claims a speaker makes, Habermas argues, do not simply reflect facts about this person's own position; they inevitably aim to be valid for all rational speakers.

This paper will be exploring Habermas's negotiation of the Kant-Hegel legacy. I will suggest that Habermas fails to appreciate the exact nature of some of the defining features of the idealist project, in particular its account of spontaneity and self-determination, which profoundly shaped Hegel's thinking. I will criticize, therefore, Habermas's claim that Hegel, despite being a protagonist of 'detranscendentalization', ultimately reverted to a 'mentalist' theory of absolute subjectivity. I will claim, moreover, that when read as a self-reflective criticism of positivity, Hegel's project can be shown to represent a challenge to Habermas's formal-pragmatic approach. In the final section, I will argue that Habermas's failure to adequately define the nature of Hegel's thinking may suggest the need for a retrieval of earlier and more manifestly Hegelian accounts of Critical Theory. At this point I will briefly invoke the work of Adorno.

Habermas's Critique of Hegel

The development of the early Hegel's thinking was profoundly shaped by his considerations of the Enlightenment, and of European modernity as such. As Habermas observes in *The Philosophical Discourse of Modernity*, the fundamental problem faced by Hegel was whether or in what sense modernity could be said to possess the conceptual and cultural resources requisite for reconciling individual aspirations to self-authorization and self-determination with objective institutional frameworks of various kinds.[6] How, indeed, is it possible to criticize and eventually overcome positivity – the arbitrary and, for Hegel, dogmatic and potentially authoritarian enforcement of normative claims and

ideals, whether in religion, art, science, philosophy, or politics? In the name of what kind of criteria or aspirations can positivity even be located and diagnosed?

In his earliest phase of serious intellectual engagement, in Tübingen, Bern and Frankfurt in the 1790s, Hegel appeared as a strong critic of the Enlightenment project, arguing that it promotes an instrumental, fragmenting, and ultimately distorted vision of rationality.[7] The Enlightenment had celebrated understanding [*Verstand*] or reflection at the expense of reason [*Vernunft*]. Thus, rather than seeking to articulate some form of absolute unity by reference to which the various dualisms of modernity – faith versus reason, theory versus practice, morality versus ethical life, the finite versus the infinite, and so on – could be overcome, the proponents of the Enlightenment merely accepted or affirmed them, thereby elevating finite, arbitrary determinations to the status of something absolute. Moreover, the main adversary of the Enlightenment, protestant orthodoxy, fared no better. By failing to transform the historical-critical activity of biblical exegesis into a living element of ethical life that could motivate moral action in accordance with reason's commands, it left the individual believer without any means to identify with religious doctrine.

Like some of his contemporaries, such as Schiller, Schelling, Hölderlin, and, a few years later, the Jena Romantics, Hegel, despite his enthusiasm for the French Revolution and for Luther's revitalization of Christianity, found modern life essentially cold and alienating; and, following the lead of Rousseau's *Second Discourse*, advocated a reconsideration of moral and cultural ideals, mainly Hellenic and Christian, borrowed from bygone epochs of human history. Unlike many intellectuals of his own generation, however, Hegel was never nostalgic about these ideals, and when searching for ways to reactualize them, he always did so within the parameters set by what he understood to be the most sophisticated manifestations of modern thought.

In his early theological writings, in which he criticizes the positivity of Christianity, Hegel turns to Kant's moral philosophy. However, from framing his account of the teachings of Jesus, in 'The Positivity of the Christian Religion', in such a way as to make Jesus be a proponent of Kant's notion of self-resolved duty, Hegel soon started to view Kantian morality as aligned with the Enlightenment trend towards greater alienation.[8] The moral law, though self-imposed, requires unconditional obedience, independently of empirical motivation and nature; yet by dividing the subject along the familiar opposition of reason and sensibility (including the capacity for passionate attachment, or what Hegel calls 'love'), it came to epitomize the fateful dialectic of externalization which, in Hegel's view, seemed to be the hallmark of Enlightenment rationalization.

Around the turn of the century, Hegel became more focused on Kant's theoretical philosophy, arguing that it was caught up in the arbitrary dogmatism of Enlightenment thinking as well. The privileging of the understanding, which offered cognitive assurance at the cost of being, as he saw it, formal, finite, and

ultimately psychologistic, over reason, the capacity to unify and reconcile the sundered 'whole', seemed unacceptable and called for a very different approach. Rather than following Kant, Hegel began developing a new systematic approach to the human mind and its relation to the social whole. Now this is the point at which Habermas starts to make his strongest claims.

Habermas's approach to the early Jena years is that '[Hegel] gambled away what, from hindsight at least, appear to be his original gains'.[9] There is, he claims, on the one hand, the 'good' Hegel of the Jena *Philosophy of Mind* and *Systementwürfe*, who was working his way toward a communicative and inter-subjectivist theory of the formation of the human mind.[10] On this 'good' account, Hegel's 'spirit', operating in the media of language and labor, is nothing but the socially instituted structures of mutual recognition that provide grounds for identity-formation, and the categories according to which the objective world is cognitively processed emerge as functions or by-products of that process. For Habermas, however, there is, also the 'bad' Hegel of the *Phenomenology of Spirit* and beyond. Habermas claims that this 'bad' Hegel replaced the detranscendentalized (or situated, finite) subject of the earlier writings with a theory of absolute subjectivity, involving a single macrosubject allegedly capable of overcoming the opposition between subjective certainty and objective sociality by reference to some form of totalizing, otherness-absorbing self-conscious 'whole'.[11] While the earlier, 'good' Hegel set the stage for thinkers such as Humboldt, Peirce, Dilthey, Dewey, Cassirer, Heidegger, and Wittgenstein, all of whom are said to have 'put the transcendental subject back into context and [situated] reason in social space and historical time',[12] the later, 'bad' Hegel returned, albeit in a grandiose and ultimately absurd fashion, philosophy to the 'mentalist' position that the earlier Hegel had found in Kant but rejected. Hegel should have remained faithful to his early view that the subject-subject model characteristic of communicative action is more fundamental than the subject-object model on which the account of alienation is predicated. He would then have been able to articulate the promises of modernity in terms of the account he provides of rational dialogue and intersubjective understanding. Instead, what he did was to reclaim the notion of free, self-reflective subjectivity, yet this time not as a finite, psychologistic entity, but as embodying the capacity for absolute freedom and hence the ability to overcome every contingency or otherness by seeing it as the product of its own self-positing.

These are large claims, and in this crude sketch I have so far said nothing about why Habermas holds them to be true. Whatever one thinks of their cogency, however, it is hard not to appreciate how well they fit in with his own rejection of 'mentalism'. As in many of his readings of other thinkers in the Western tradition, there is here a promise that has been rejected, forgotten or repressed; thus by offering his dual interpretation, Habermas can place himself at the end of a history of failures and tell his audience exactly what went wrong.

For the moment, I will neither deal with Habermas's own theory of communicative rationality, nor with the corresponding, though rudimentary, intersubjectivist account he claims to have found in Hegel's Jena *Systementwürfe*. What interests me, rather, are the views he brings to bear on the notion of spirit and the role this entity is supposed to play in Hegel's version of the 'self-grounding of modernity'. Is it plausible that Hegel's theory of freely self-determining spirit, as it is developed and articulated in the *Phenomenology of Spirit*, should be viewed as a form of mentalism?

By 'mentalism', Habermas refers to any philosophical theory that, following the lead of Descartes's epistemology, takes *the* crucial challenge of philosophy to consist in the problem of accounting for the possibility of objective representation. A mentalist is someone who believes that objective representation (if at all possible given the skeptical doubts which, in Descartes, were epitomized in the construction of the famous dream-argument) takes place through some sort of ideational 'mirroring', and that such mirroring – available through introspection – can only occur for a self-conscious epistemic subject. As opposed to the, at best, indirect access one may have to mind-external or mind-independent reality, epistemic access to one's own private mental content is on the mentalist view understood to be immediate and incorrigible. Moreover, since there is no immediate access to a mind-independent reality, truth and objectivity become functions of the quality or mode with which ideas (or representational content in general) are presented to the mind. On the assumption that the mind actually is able to 'mirror' reality, Descartes famously introduced the criteria of clarity and distinctness in order to distinguish veridical from non-veridical representations. In acts of self-reflection, the subject assesses the degree of objectivity of its own ideas.

One difficulty which arises when assessing Habermas's lumping of both Kant and (the 'bad') Hegel with Descartes and mentalism is that it seems to do little justice to the specific form of idealism which emerges in Germany with the publication of the *Critique of Pure Reason*, and which, though in a radically reshaped fashion, is being continued in Hegel's *Phenomenology*.[13] The Cartesian (mentalist) view presupposes a robust realism according to which the world to be correctly represented is the world as it is independently of a subject's conceptual determination of it. Given a non-inferentially warranted mental state, the mentalist asks, how can we prove whether it corresponds to a mind-independent reality? Yet neither Kant nor Hegel believe a) that there is such a thing as a world *in itself* to be represented at all, or b) that the most promising way to reconstruct the conditions under which knowledge is possible is to try to identify an immediately given mental state and then inquire into its possible correspondence with a mind-independent object. Central to both Kant's and Hegel's projects is that non-inferentially warranted states cannot have an epistemic value because in order to take a mental state to possess any kind of determinacy (and it must have determinacy, be of something in particular, in order to represent something), it is necessary to take it to have a specific content,

and one can only do so by making a judgment about it that involves the use of concepts. There is, on Kant's and Hegel's views, no intuitive knowledge of anything; all knowledge requires the mediation of concepts deployed in judgments. Rather than taking the mind's epistemic role to consist in passively representing the world, where the truth-claim of a given representation is assessed in terms of specific epistemic qualities, both Kant and Hegel argue that the mind is fundamentally active in that, in getting to know how things stand, it determines what is given to it by relating judgmentally (and hence apperceptively) to the given. To make judgments, then, and thereby to experience mind-independent objects, is for Kant and Hegel to submit oneself to the specific norms that govern their formation.[14] It is only when we actively process what is given by placing it judgmentally in normatively structured relations that questions of objectivity, and therefore of truth and falsehood, can emerge.

For Kant, such considerations achieve further elaboration when reflecting upon what it means to take representations to be representing an object. I am able to view myself as apperceptively self-aware of my representations *qua* representations of an object insofar as I take them to be combined and determined in certain ways that correspond to the possible forms of judgment. Whatever one's views are concerning Kant's influence on Hegel, the Hegel of the *Phenomenology* accepts that judging, or making a claim to knowledge, is an activity that presupposes a pre-given commitment to specific norms that determine what counts as authoritative or objective for the kind of self-conscious experience of objects that the idealists take to be required for there being a genuinely cognitive relation to the world. Yet on his account there is no fixed and transhistorical set of 'forms of judgment' such as Kant's categories that govern all possible judging; rather, the *Phenomenology* is a progressive testing of successive candidates for successful notional determination of judgment. For each putative experience, Hegel presents a test-case which will show whether the notional determination is adequate for the self-conscious experience of objects that is being promised. If it is not adequate (which it is bound to be, given Hegel's desire to recount a developmental story that will gradually lead to greater insight into the subject's own free responsibility for the notional determinations it applies), offering conflict, incoherence and paradox rather than self-conscious experience of objects, then a new notional determination is introduced which resolves the problems and disappointments of the former. In the final instance, the author of the *Phenomenology* will be arguing that the ultimate basis for epistemic authority is a community of agents – an aspect of what Hegel calls 'spirit' [*Geist*] – in which every member is recognized by the others as free and equal, and in which the members take their self-reflective form of life to be the absolute vantage-point from which any question of epistemic authority can be raised.[15] At no point, then, is Hegel venturing beyond the level of consciousness's own dialectical self-reflection in order to assess its norms with reference to the object as it is in itself. The only mentalism at stake in the *Phenomenology* is the one which Hegel himself introduces in the

initial chapter on 'consciousness' only to discard it by showing that its claims about what counts as knowledge generate inconsistencies and paradoxes.[16]

In a recent essay on Kant's and Hegel's epistemologies, 'From Kant to Hegel and Back Again: The Move toward Detranscendentalization', Habermas recognizes the need to distinguish, within the group of positions which he characterizes as mentalist, between Cartesian realism and the transcendental turn in Kant and subsequent idealists.[17] The transcendental turn involves, he writes, 'the idea that the knowing subject determines the conditions under which it can be affected by sensory input'.[18] Yet the way in which Kant's project is subsequently described reveals why he continues to think of it as wedded to mentalism. Kant, he claims, 'wishes to solve a problem that he inherits from the mentalist paradigm, one that establishes the contrast between a representing subject and a world of objects offered for representation. At the same time, he also inherits those unanalyzed notions of subjectivity and self-reflection that are constitutive for the mentalist framework'.[19] This interpretation, however, is misleading. By understanding the central problem of Kantian epistemology to reside in the quest for correct representation of 'a world of objects offered for representation', Habermas fails to realize that the crucial issue in Kant is not representation but the uncovering of the *conditions* our representations must conform to in order for human agents to be able to take them as representing objects in the first place. The second claim, about subjectivity and self-reflection, is not easy to understand. However, what Habermas seems to get at is that Kant's conception of the transcendental apperception involves what he calls 'a self-reflection that operates as a representation of my own representings'.[20] Again, this is unpromising as an account of Kant's position. If self-consciousness (or self-reflection) were thought to be some sort of second-order representation, then that would require yet another level of self-awareness again, and so on ad infinitum. We would, as Dieter Henrich has argued, be faced with an infinite regress.[21] However, as I have already indicated, Kant is not committed to such a view. Very roughly, for him the transcendental unity of apperception is the capacity to take oneself as the author of one's experience, and therefore of the judgments made in relation to it. The subject's identity is a function of the capacity to ascribe experiences to one and the same self over time, which ultimately is made possible by the way in which the same subject is able to create a synthetic unity among its experiences.

Habermas's recent ascription of mentalism to Hegel is more complex. He does acknowledge that the specific form of self-reflection being practiced in the *Phenomenology* does not fall immediately under the mentalist paradigm. In clear opposition to Kant's transcendental epistemology, it involves:

a) a learning process whereby the subject comes to realize that its self-reflectively accepted notion of what counts as knowledge is ultimately dependent on standards that are upheld in and through communal practices of mutual recognition; and

b) that the orientation of the reconstructive philosopher therefore needs to change from that of being preoccupied with the spontaneity of a solitary transcendental subject to what eventually, when the Hegelian narrative reaches the level of 'Spirit', turns into an account of the intersubjective constitution of the objective world.

One might add to this that the *Phenomenology* can be read as one continuous battle with Cartesian realism. Starting, in the first chapter, with the experience of an individual consciousness embodying strong realist claims about knowledge, Hegel gradually undermines this picture by showing how experience is mediated – first, through conceptual determinations as such, and second, through communal recognition of the experiencing subject's conceptual determinations.

Why, then, can Hegel continue to be accused of mentalism? Habermas's argument is this. The intersubjectivist reading of spirit, while attractive in that it overcomes the 'methodological solipsism' of the Descartes–Kant–Fichte tradition, is 'deficient by Hegel's own standards'.[22] It is deficient because the intersubjectivist reading fails to bridge the gap between claims to knowledge that satisfy intersubjective criteria of validation and what is objectively true. It is simply not the case that idealized intersubjective validation implies objective truth. Thus, even if a contestable view has turned out to be acceptable 'for us', it has still not been established that it is true of the world as it is independently of the intersubjectively established framework. On Habermas's account, Hegel would see the intersubjectivist reading of spirit as involving a form of arbitrariness with regard to the norms which govern the community's reason-giving activities.

According to Habermas's *own* thinking, such arbitrariness must be accepted. It is simply another way of characterizing the postmetaphysical fact that the most authoritative source of knowledge in modern societies, namely science, can only develop in a spirit of inevitable fallibility. Scientific research is finite: it must humbly accept that what today counts as a rationally acceptable theory may tomorrow be overthrown by one of its competitors. However, in what Kuhnian philosophy of science has taught us, it is required that the scientist is open not only to piece-meal revisions, but also to anomalies that, if impossible to integrate into the adopted framework, may ultimately lead to a change of paradigm. This is another aspect of the inevitable arbitrariness characterizing all knowledge. It should be noted, though, that Habermas at this point is conflating two different claims. It is one thing to say that those who engage in scientific research must be open to the possibility of critique, revision and falsification, and that that involves the acceptance of a certain form of arbitrariness; it is quite a different thing, however, to say that the framework within which such research takes place – the normatively structured relationships of inference that must be in place for data to be interpreted and theories to be established and corroborated – is arbitrary. For us to be able to think of it as

arbitrary it would be necessary to envision alternatives that would be equally well grounded in the learning process of a given culture, yet such alternatives are precisely not available for the simple reason that we could not make sense of them within our own life-form. It also suggests, along with the metaphysical realism which Habermas elsewhere wants to reject, that there is such a thing as the world as it is in itself, and that cultural standards must be arbitrary because they do not present us with an absolute view-point. Yet one of the central lessons of Hegel's thinking is precisely that the very idea of the world as it is in itself is incoherent: there is no objecthood independently of the conceptual scheme we employ. Only if we saw the conceptual scheme as imposed upon a world that somehow were epistemically constituted outside of our concepts would it make sense to ask whether it accurately matches or represents the world. However, the Hegelian view is that no such imposition takes place. Our rules for determining objects stretch out all the way to the object; thus discovering empirically what the world is like will not cast any light whatsoever on the constitutive or transcendental relation between thought and world.[23] Only the examination, as in the *Phenomenology*, of the different ways we can *take* the world to be determinate will reveal this relation.

Regardless of how Habermas would go about defending this particular version of contextualism (which certainly does suggest, along the lines of the mentalist paradigm, that there is a gap between our cognitive practices and the world as it is in itself), the view he attributes to Hegel is that the intersubjectivist understanding of objective spirit is insufficient to ground objectivity in the way Hegel (on this construal) would seem to want it to do. Instead, what is required (according to Habermas's Hegel) is an account of *absolute* spirit in which an 'absolute subject' is posited as the basis from which the history of consciousness emerges: 'This subject is thought of as the One and All, as the totality that 'can have nothing outside itself".'[24]

The account Habermas invokes here is a version of Platonism whereby spirit, or the absolute subject, is considered to be a metaphysical entity – or, more correctly, a noumenal reality – that actualizes itself teleologically by progressively overcoming its own self-alienations. In analogy with Fichte's *Tathandlung* (or self-positing ego) yet expanded beyond finite consciousness, spirit becomes a 'self-reflection writ large'[25] that permeates the human mind, world history, and external nature. While this represents Hegel's solution to the arbitrariness-problem in that noumenal reality is now accessible for a philosophy of spirit, it can at the same time be interpreted as mentalism taken to its utmost extreme: for in contrast to the more modest conceptions of subjectivity that we find in Descartes or Kant, the Hegelian absolute subject expands to cosmic dimensions and becomes the source of all reality. Knowing the way it determines itself throughout its formative processes becomes equivalent with possessing the kind of absolute knowledge which Hegel's mature system promises.

Habermas's interpretation of Hegel may sound excessive. However, as Frederick Beiser points out, not only has it been pervasive in much scholarship on

German idealism but, insofar as Hegel is considered to represent the culmination of the Cartesian tradition, enjoys a certain inherent plausibility.[26] If the Cartesian (or mentalist) is committed to the view that the subject has immediate knowledge only of its own ideas, and therefore that it has no direct recourse to that which is beyond its circle of awareness, then the German objective idealists attempt to avoid the skeptical conclusion arising from this paradigm by widening consciousness to embrace everything: thus knowing the expanded mind's own ideas, which amounts to self-knowledge, becomes equivalent with having metaphysical knowledge that extends into the noumenal realm. By combining this idea with Kant's alleged notion of the world-creative powers of the subject, one could reach the kind of position that Habermas attributes to Hegel.

According to Beiser, however, this story has little or no truth to it:

> the absolute subject, the infinite ego, or universal spirit, *understood as a metaphysical principle or noumenal reality*, never had much of a role to play in German idealism. If it appeared at all, it was only very briefly, confined to a very short phase of Schelling's philosophical project, the few months he adhered to the doctrines espoused in his early *Vom Ich als Prinzip der Philosophie* (1796); and Schelling quickly moved away from this position, and even during this period he equivocated whether he meant to commit himself to the existence of the absolute subject.[27]

Beiser's own account, however, is different from the one I already outlined with regard to Hegel. According to Beiser, the German idealists, including Hölderlin, Schlegel, Novalis, Schelling, and Hegel, instead turned to 'something impersonal, neutral, or indifferent, whether it be pure being, life, or the indifference point; to construe it as the ego is to hypostatize and anthropomorphize it, dragging it down into the realm of finite experience'.[28]

I do not intend to discuss Beiser's positive account but will simply note that if Habermas's interpretive hypothesis were correct, then the question would arise of how finite humans may have epistemic access to the realm of absolute or infinite spirit.[29] Hegel would then fall behind Kant's critique of rationalism and introduce a version of transcendental realism that would be indefensible on the basis of the conceptual resources on offer in the *Phenomenology*, let alone on independent grounds. Habermas's claim that the adoption of an extremely ambitious metaphysical theory of absolute spirit would make it easier for Hegel to defend the unconditional objectivity of what is rationally acceptable 'for us' falters in that it so blatantly disregards that so much of the philosophical labor being performed in the *Phenomenology* is geared towards undermining the very possibility of, and rationale for, asking whether our conceptual schemas, or what rational beings can accept as authoritative, actually corresponds in some 'deep' sense to what is 'really' out there. If Habermas's

reading were right, then Hegel, rather than succeeding in establishing transcendental conditions for objective judging, would commit himself to a view that would generate a radical form of skepticism.

Habermas's Kantianism

Much of Habermas's effort to establish a theory of rationality that, when added to his account of 'systemic intervention in the lifeworld',[30] can provide a more balanced understanding of modernity than those of both Hegel and most of the European thinkers who responded to his work over the next 120 years or so is predicated on a return to motives found in Kant.[31] We have seen that Habermas (wrongly in my view) believes to have located problems in Kant's theoretical philosophy that are inherited from the Cartesian or mentalist tradition; thus, returning to Kant must involve finding an alternative Kant who is *not* committed to, or in any sense dependent on, a mentalist framework. In Habermas's formulation of the task, it is necessary to look for 'paradigm-neutral types of self-reflection' in Kant's writings. Thus, in the important formulation of what he calls a formal pragmatics, he seeks to reconstruct the pragmatic presuppositions speakers allegedly must make when engaging in 'action oriented towards understanding' [*verständigungsorientertes Handeln*].[32] According to this theory, speakers understand and interpret each other in light of the reasons being offered for their respective claims, and when being oriented exclusively towards reasons they will inevitably have to take each other to be free and equal in senses Habermas spells out in great detail. In particular, Habermas introduces the notion of a specific set of idealizations that are said to underlie everyday speech and make up necessary and universal commitments that every rational speaker must undertake when communicating and discussing. Such idealizations, or 'discourse rules', stipulate for example that participants in discourse must have the same chance to put forward or call into question claims that are being raised, and that the discussion is free from distorting influences, whether their source is open domination, conscious strategic behavior, or the more subtle barriers to communication deriving from self-deception. Drawing on some of the deepest ideals of the German idealist tradition, what this means is that linguistic activity commits us to the assumption that other participants in the linguistic community are to be understood as fellow subjects, not manipulable objects, and so in all cases as potentially rational subjects, freely capable of advancing or rejecting various claims being made by subjects on each other. In his discourse ethics, Habermas even attempts to derive a moral principle from these rules, arguing that every serious speaker is committed to this principle when engaging in moral debate over disputed social norms.[33]

Yet what is the exact theoretical status of these rules? According to Habermas's close associate over many years, Karl-Otto Apel, the discourse rules – which, as outlining an indefinite community of mutually recognizing speakers,

for him anticipate an ideal form of life – are of a transcendental nature: they articulate the apodictical requirements that every speaker must accept as binding.[34] They pretend, we might say, to be universally valid in roughly the same sense as Kant's categories are meant to be so. For Habermas, however, the formulation of the rules is itself hypothetical, the result of a collaboration between the empirical efforts of social and linguistic sciences, and the philosophical efforts of rational reconstruction.[35] One problem with this weaker account, though, is that it is not clear what would count as an empirically based disclaimer of the formulation. Is not a person who formulates an alternative set also committed to the discourse rules, and how can she then present her revision without presupposing that which she is about to criticize? Moreover, if the discourse rules really just have an empirical status, then how can Habermas draw on them in order to formulate a view of rationality that, as he wants, can withstand relativist skepticism? It is not enough to suggest that no one can engage, say, in rational debate without accepting the idealizations that the discourse rules stipulate if other cultures turn out to possess different language games for adjudicating controversial claims and proposals.[36] In short, if the discourse rules can be supported or objected to in light of empirical considerations, then they cannot ground what Habermas wants them to establish, namely a form of strong universalism. Habermas cannot have both: a strong universalism *and* an empirically based reconstruction of rules of discourse.

Another and possibly more interesting problem is that discourses may seem to satisfy the ideal requirements which Habermas outlines and still not deserve to be called rational. In evaluating the rationality of specific discourses, there will always be a need for context-sensitive and historically specific judgment. Why, one may for example ask, is it always rational to grant every participant in a rational debate the same right to intervene and raise objections, etc. when it seems widely, if not universally, accepted, both in theoretical and practical contexts, that some people will always be more competent, experienced, and insightful than others? To say that their competence, experience and insight will have to be proven in the discourse is a non-starter: if so, then very few serious discussions, whether in science or philosophy, would ever get started. No rational discourse can proceed without exclusions, yet how, when and the degree to which such exclusions should be carried out will necessarily be a matter of judgment. Needless to say, discourses take place in real life – that is, in unruly and singular situations that require decisions of various kinds.

At this point it may be useful to remind oneself of Hegel's critique of positivity. After all, one of the central claims that Hegel makes in this regard is that the idea of 'first philosophy' that informs a Kantian transcendental critique (and, by implication, Habermas's formal pragmatics, even though he seeks to soften it by introducing empirical constraints) is deeply problematic. There is always something prior and historically given on which reflection depends, hence the constitution of epistemic norms cannot be construed as an

absolute origin but must itself be the outcome of historically mediated processes of self-reflection and immanent critique. However, by contrast to the position Habermas has been defending since the mid-1970s, which seeks to combine empirical and transcendental reflection in an external relation, Hegel can agree that claims concerning fundamental structures of rationality are themselves of a historical nature *without* accepting that they therefore must be possible to criticize by reference to empirical considerations. Habermas concedes this point in his 1968 discussion of Hegel in *Knowledge and Human Interests* but ignores it as soon as he starts developing his formal pragmatics.[37] The necessities which each formation of consciousness in the *Phenomenology* take as given are necessities – necessary epistemic norms – in the sense that they stake out what a particular formation of consciousness takes to be determinative for itself. They are so fundamental that it would make no sense either to support or disclaim them empirically. The claim that knowledge, as in the model of 'sense certainty' with which Hegel sparks off his dialectic, is equivalent with whatever presents itself immediately to consciousness, cannot be touched by empirical considerations: it can neither be supported nor falsified by such claims. The only procedure that can offer a rational testing of such a claim is the one which Hegel himself proposes, namely an immanent critique that seeks to verify whether the claim (or norm) *on its own terms* offers the kind of epistemic achievement that it promises. Does the immediacy of whatever is presented to consciousness really qualify as knowledge in the sense outlined by this formation of consciousness as it starts its process of self-reflection? As it turns out, it will experience failure: there can be no claim to knowledge unless the content of the claim has some form of determinacy, and Hegel will show that such determinacy is only possible insofar as concepts are being used to individuate and discriminate the given.[38]

I am not proposing that Hegel's *Phenomenology of Spirit* and its critique of positivity can be directly reclaimed as an account of rationality today. Hegel's ambition, for example, of showing that there is a necessary relation between each formation of consciousness, such that given the properties of formation X, a specific and unique formation Y will necessarily follow, introduces external demands (in Hegel's case the logical-dialectical entailment-relations explored in *The Science of Logic*) on the process of self-reflection that are at odds with the claim to autonomy ascribed to each formation.[39] If such autonomy is granted, then the outcome of the dialectical self-reflection may simply be whatever determination there is that resolves the epistemic quandaries of the original formation. If formation Y follows from formation X, then it is not because it is the *only* formation that satisfies the requirements arising from the breakdown of formation X, but simply because it satisfies them. Y is therefore a possible but not necessary consequence of X.[40]

Immanent critique takes the existence of language games and human practices as given and, without any attempt to establish foundations or presuppositionless beginnings, engages in critical self-reflection with a view to

showing whether they offer what they purport to offer according to their criteria and conceptual determinations. Since the criteria (of what counts as something in particular, or of whether something is valid or right) and conceptual determinations make up the historical framework within which human speech and activity can appear as intelligible, they do not have an empirical status, nor can empirical considerations be appealed to in the process of reflecting upon them.

In *Knowledge and Human Interests* Habermas formulates a similar view.[41] Drawing on Fichte's critique of dogmatism, Habermas argues that there exists a form of self-reflection which is characterized precisely by the desire to achieve mature autonomy through the overcoming of false hypostatizations. As in Fichte, reason is not, he argues, a theoretical faculty, a faculty of contemplation, so much as it is the unification of theoretical and practical interest in the search for one's own autonomy; and Habermas's goal in this book is to reinstate this model as essential to the very endeavor of critical social theory in the Frankfurt School tradition:

> For reflection destroys, along with a false view of things, the dogmatic attitudes of a habitual form of life. . . . In false consciousness, knowing and willing are still joined. . . . The reversal of consciousness means the dissolution of identifications, the breaking of fixations, and the destruction of projections.[42]

In the 'Postscript' to *Knowledge and Human Interests*, however, Habermas, as already mentioned, starts to distance himself from this notion of *critique*. In particular, he finds that it harbors an equivocation between two very different notions, namely 'immanent critique' as already outlined and 'rational reconstruction' in the Kantian sense.[43] Of course, detecting such an equivocation in the historical use of a term does nothing to show that there actually exist philosophical *reasons* to add a program of rational reconstruction to the initially conceived program of immanent critique, and it certainly does not license the gradual abandonment of the latter in favor of the former. Perhaps the most weighty reason for differentiating between the two forms of critique and developing a theory of formal pragmatics and rationality was that, on Habermas's later view, it no longer seemed clear that the unmasking of ideological hypostatizations and world views had any intersubjectively valid standard to appeal to in characterizing perceptions and theories as distorted or false. Having abandoned what he thought of as Hegel's notion of absolute truth or the absolute subject (as articulated, ultimately, in *The Science of Logic*), it seemed to him that no alternative existed but to construct a theory that would account for the discursive commitments necessary in order to raise claims to intersubjective validity. Such claims are then meant to be redeemed in discourses that are rational according to the stipulations of the theory.

An earlier account of critical theory

We have seen that not only are there reasons to doubt the validity of Habermas's interpretation of Hegel, but arguments drawn from Hegel may well be used in order to question Habermas's own position. I will now complicate this picture considerably (and possibly too much) by introducing the thought that resources for obtaining a more refined conception of what immanent critique involves may perhaps exist not only in the early Habermas (who, as I said, quickly abandoned this procedure, or at least downplayed its theoretical role), but in earlier versions of Critical Theory, in particular those that are more directly influenced by Hegel than Habermas has been. One such early version is found in Adorno's negative dialectics. Like Habermas, Adorno does launch a high-pitched critique of Hegel, accusing him of adopting a pre-critical metaphysics of the absolute subject along the lines already explored. According to Adorno, Hegel's system falsely reduces everything to identity; it is, he writes, the 'belly turned mind', an extreme idealism that is incapable of respecting any theory- or mind-independent evidence.[44] Yet while making these accusations, however dubitable, Adorno nevertheless positions himself as a distinctly Hegelian theorist. He does so, I will claim, by conceiving of philosophy as a form of radical self-reflection capable of challenging the prevailing, more instrumentalist forms of reason in modernity.

The notion of instrumental reason – essentially end-indifferent, procedural and technocratic reasoning – is central to the early Frankfurt School's assessment of modernity but gets dramatized by Adorno to become the driving motor of human history as such. In his most philosophical writings, Adorno sees instrumental reason and the forms of domination on which its exercise is based as totalitarian and deeply damaging of both interhuman relations as well as the relation between man and nature.

It has occasionally been claimed, not least by Habermas himself, that Adorno's dark account of instrumental reason and its pervasiveness in late modernity undermines his program of self-reflection and critique.[45] If reason is instrumental and only capable of assessing the best means to given and therefore, from the standpoint of reason, arbitrary ends, then the critical theorist is sawing off the branch on which he sits: his practice of critique is then performatively at odds with his theory of rationality, and he will have to accept that he does not dispose of any normative resources on which to base his critique. Now Habermas's point would certainly be well-taken if it were indeed true that Adorno *needs* an independent theory of rationality to back up his claims. However, he would only need that if he accepted Habermas's anti-Hegelian premise, namely that a critique of life-forms and forms of rationality must be supported by a quasi-transcendental, universalist theory of rationality. The position we find in Adorno is rather that critique must proceed on an immanent basis. Although a theorist must always be open to the possibility of radical conceptual revision, there is no coherent standpoint beyond the practices that

already exist. In his *Negative Dialectics*, Adorno tries to develop this point by arguing that the task of philosophically and socially motivated critique is to criticize the failure of specific norms or criteria to deliver the epistemic results they promise. On Adorno's account, the most pressing example of such a critique is one that focuses on 'identity', a notion which is best and most fruitfully interpreted in terms of Hegel's theme of positivity. This, at least, is how I read passages such as the following:

> The name of dialectics says no more, to begin with, than that objects do not go into their concepts without leaving a remainder, that they come to contradict the traditional norm of adequacy. ... It indicates the untruth of identity, the fact that the concept does not exhaust the thing conceived. Aware that the conceptual totality is mere appearance, I have no way but to break immanently, in its own measure, through the appearance of total identity.[46]

What Adorno claims here is that knowledge of objects is made possible by norms, concepts or rules – that is, the criteria we dispose of for knowing the object in its objecthood. When they fail to yield such knowledge, an immanent process of conceptual revision is set in motion.

It would have served his purpose better had Adorno made it clear that the target of his critique is *false* identity and not identity per se. False identity, for Adorno, arises when epistemic norms or criteria are naturalized and viewed as expressions of an immutable order of things in themselves, beyond the contingencies of history. Such naturalized norms or criteria are thus claimed to provide an 'absolute identity' between knowledge and its object, or, in Adorno's Hegelian jargon, between subject and object. On Adorno's Weberian view of rationalization, the norms and criteria we possess as modern agents tend increasingly to provide some kind of higher-order abstraction in accordance with which the object of knowledge is supposed to be understood. Natural science, for example, typically presents its claims in terms of some formal mathematical procedure, and knowing the world becomes a matter of knowing it in those terms.[47] Likewise, the phenomenon which Adorno, following Marx, calls 'commodification' (and which can be traced back to Hegel's preoccupation with dogmatism) involves a reduction or transformation of the object (the commodified entity) to the terms and conditions provided by a capitalist system of exchange. In both cases – the Galilean ontology of natural science and the system of commodities in a modern market economy – one may argue that norms or criteria of what counts as knowing something can appear to call for further reflection. Adorno is not simply proposing that there is something inherently flawed in the way science and the economy operate; rather, he is inviting the critical theorist to consider whether what we take ourselves to know about certain objects – the physical world, the world of merchandise – really is fully coherent. Does science *on its own terms* offer the

kind of knowledge it promises, or is perhaps quantification somehow at odds with the ambition, say, to also understand singular (and hence non-repeatable) processes and events? Does the notion of the commodity, of things being defined in terms of their exchange value, exhaust what objects are for us?

We do not need to accept what I have called Adorno's Weberian pre-suppositions in order to see that the notion of immanent critique can form the basis for an ambitious program of philosophical self-reflection.[48] On Adorno's view, philosophy is a particular form of self-reflexive critique that ultimately aims to place the individual – the subject – in a position from which normative commitments can rightfully be viewed as a rationally self-chosen result of processes of reflection. At the same time, however, Adorno is deeply skeptical about the very possibility of achieving such a position. On his account, which radicalizes Hegel's understanding of positivity, modernity, with its dominance of formal and instrumentalized modes of reasoning, offers very few spaces in which to exercise such a capacity. Moreover, to the extent that they do exist – in philosophy, but also in certain types of responses to the advanced modernist arts – he worries that the exercise of reflection, if taken to be capable of arriving at a reconciliation between the criteria of knowledge and the claims to it, could lapse into an unjustified affirmation of the given. Dialectical thinking should mainly keep the possibility of resolution and reconciliation open. It should focus on the incompatibility of claim and criterion, or, as Hegel puts it, subjective certainty and truth, yet avoid thinking that the progressive move towards a new formation that will reconcile specific claims to knowledge with their epistemic presuppositions can be performed in abstraction from its accompanying social conditions.[49] For Adorno, critique and social change go together: there cannot be an immanent critique that is not at the same time demanding social change.

We should now be able to see that despite Adorno's commitment to imma-nent critique in a roughly Hegelian sense, there are some crucial differences between his and Hegel's understanding of dialectics. Adorno avails himself of dialectical self-reflection in a much more provisional manner than Hegel. He is constantly concerned to avoid the temptations of systematicity and complete-ness. However, the claim to inherit the Hegelian project for the purpose of articulating a critical theory of modernity does carry considerable force and should be seen as offering a worthy competitor to Habermas's neo-Kantianism.

Notes

1 For a useful introduction to what such a Hegelian notion of critique involves, see Garbis Kortian, *Métacritique* (Paris: Les Editions de Minuit, 1979), pp. 24–38.
2 Perhaps the most programmatic and influential statement of this position within the history of the Frankfurt School is Max Horkheimer's 1937 essay 'Traditional and Critical Theory', in *Critical Theory*, trans. M.J.O. O'Connell *et al.* (New York: Herder & Herder, 1972), pp. 188–214. For another and equally important attempt to articulate the 'logic' of immanent critique, see Herbert Marcuse, *Reason and Revolution:*

Hegel and the Rise of Social Theory (London: Routledge, 1968), esp. pp. 16–29. In addition to Hegel, the 'classical' exercise of this type of critique is Marx's analysis and exposition of the internal contradictions of political economy. When the implications of categories such as labor and value are thought through to the end, they turn out to have a meaning which is incompatible with what they have when being applied in everyday, pre-theoretical practice.

3 By 'the early Marx' I mean in particular the author of the 'Introduction' to the 1844 *Contribution to the Critique of Hegel's Philosophy of Right* and the *Economico-Philosophical Manuscripts* of the same year. Both are printed in *The Portable Karl Marx*, ed. Eugene Kamenka (London/New York: Penguin, 1983), pp. 115–25 and pp. 131–52.

4 Jürgen Habermas, *Knowledge and Human Interests*, trans. Jeremy Shapiro (Cambridge: Polity Press, 1987), pp. 379 ff.

5 For a particularly succinct discussion of the difference between Habermas's Hegelian and Kantian side when it comes to issues of rationality, see Richard Rorty's review of Habermas's *Truth and Justification*, trans. Barbara Fultner (Cambridge, Mass.: The MIT Press, 2003) in *Notre Dame Philosophical Reviews* (12 August 2003).

6 Jürgen Habermas, *The Philosophical Discourse of Modernity*, trans. Frederick Lawrence (Cambridge, Mass.: The MIT Press, 1987), p. 25: 'As it seemed to the young Hegel, a *positivity of ethical life* was the signature of the age'. Habermas (p. 27) adds that 'In these years around 1800, Hegel made a case for the verdict that both – religion and state – had degenerated into sheer mechanisms, into a clockwork, into a machine'. The relevant texts by Hegel are *The Difference between Fichte's and Schelling's System of Philosophy*, trans. H.S. Harris and W. Cerf (Albany, N.Y.: SUNY Press, 1977); *Faith and Knowledge*, trans. H.S. Harris and W. Cerf (Albany, N.Y.: SUNY Press, 1977); and *Early Theological Writings*, trans. T.M. Knox and R. Kroner (Chicago, Ill.: University of Chicago Press, 1948).

7 For a good overview of Hegel's development in this period, see H.S. Harris, 'Hegel's Development to 1807', in Frederick C. Beiser (ed.), *The Cambridge Companion to Hegel* (Cambridge: Cambridge University Press, 1993), pp. 25–51. See also Terry Pinkard, *Hegel: A Biography* (Cambridge: Cambridge University Press, 2000), pp. 19–117.

8 Hegel, *Early Theological Writings*, pp. 67–181. In 'The Spirit of Christianity and Its Fate' (ibid., pp. 182–301), Hegel is much more critical of Kantian morality, arguing that it collapses into sheer legality – that is, into something alien.

9 Jürgen Habermas, 'From Kant to Hegel and Back Again: The Move toward Detranscendentalization', in *Truth and Justification*, trans. Barbara Fultner (Cambridge, Mass.: The MIT Press, 2003) p. 176. Habermas's works contain numerous discussions of, and references to, Hegel. The account offered in *Truth and Justification* can be read as a slightly revised restatement of claims made in the following three, previously published, texts: 'Labor and Interaction: Remarks on Hegel's Jena *Philosophy of Mind*', in *Theory and Practice*, trans. John Viertel (Boston: Beacon Press, 1973), pp. 142–69; *Knowledge and Human Interests*, trans. Jeremy Shapiro (Boston: Beacon Press, 1971), pp. 3–24; *The Philosophical Discourse of Modernity*, pp. 23–44.

10 Habermas is here drawing on Hegel, *Jenaer Systementwürfe* I-III (Hamburg: Felix Meiner, 1986).

11 For alternative interpretations of Hegel's development in Jena, see Rolf-Peter Horstmann, 'Probleme der Wandlung in Hegels Jenaer Systemkonzeption', *Philosophischer Rundschau* 19 (1972), pp. 87–118; Heinz Kimmerle, *Das Problem der Abgeschlossenheit des Denkens. Hegels System der Philosophie in den Jahren 1800–04* (Bonn: Bouvier, 1970); H.S. Harris, *Hegel's Development: Night Thoughts* (Oxford: Clarendon Press, 1983); Klaus Düsing, *Das Problem der Subjektivität in Hegels Logik* (Bonn: Bouver, 1976); Otto Pöggeler, 'Hegels Jenaer Systemkonzeption', in *Hegels Idee einer Phänomenologie des Geistes* (Freiburg: Verlag Karl Alber, 1973); and Robert Pippin,

Hegel's Idealism: The Satisfactions of Self-Consciousness (Cambridge: Cambridge University Press, 1989).

12 Habermas, *Truth and Justification*, p. 175.

13 For a succinct statement to the effect that Descartes, Kant and Hegel were united in such a common cause, see Habermas's discussion of 'prima philosophia as philosophy of consciousness' in *Postmetaphysical Thinking*, trans. William Mark Hohengarten (Cambridge, Mass.: The MIT Press, 1992), p. 31:

> Self-consciousness, the relationship of the knowing subject to itself, has since Descartes offered the key to the inner and absolutely certain sphere of the representations we have of objects. Thus, in German Idealism metaphysical thinking could take the form of theories of subjectivity. Either self-consciousness is put into a foundational position as the spontaneous source of transcendental accomplishments, or as spirit it is itself elevated to the position of the absolute. The ideal essences are transformed into the categorial determinations of a productive reason, so that in a peculiarly reflexive turn everything is now related to the one of a generative subjectivity. Whether reason is now approached in *foundationalist* terms as a subjectivity that makes possible the world as a whole, or whether it is conceived *dialectically* as a spirit that recovers itself in a procession through nature and history, in either case reason is active as a simultaneously totalizing and self-referential reflection.

What I say about the lumping of Descartes, Kant and Hegel together is indebted to Robert Pippin's *Hegel's Idealism*, pp. 3–41, and to his important article 'Hegel, Modernity, and Habermas', in *Idealism as Modernism: Hegelian Variations* (Cambridge/New York: Cambridge University Press, 1997), esp. pp. 161–63.

14 This is a dominating claim in Kant's first Critique. See for example *Critique of Pure Reason*, trans. Norman Kemp Smith (London: MacMillan, 1986), B102:

> Transcendental logic (. . .) has lying before it a manifold of *a priori* sensibility, presented by transcendental aesthetic, as material for the concepts of pure understanding. (. . .) But if this manifold is to be known, the spontaneity of our thought requires that it be gone through in a certain way, taken up, and connected. This act I name *synthesis*.

In Hegel's *Phenomenology of Spirit*, trans. A.V. Miller (Oxford: Oxford University Press, 1977), this idea enters explicitly at the end of the dialectic of consciousness – that is, at the end of the first chapter.

15 When I say that the notion of a self-reflective, rational community is an 'aspect' of what Hegel means by *Geist*, I mean to hold the door open to the view that *Geist also*, for Hegel, has a divine status. Without being able to show this here, I believe that Hegel predominantly held the view that *Geist* could be both the self-reflective rational community *and* God. For some useful reflections on this issue, see Jean Hyppolite, *Genesis and Structure of Hegel's Phenomenology of Spirit*, trans. Samuel Cherniak and John Heckman (Evanston: Northwestern University Press, 1974), pp. 29–31.

16 Hegel, *Phenomenology of Spirit*, pp. 58–103.

17 Habermas, *Truth and Justification*, ch. 4.

18 Ibid., p. 179.

19 Ibid., p. 180.

20 Ibid.

21 Dieter Henrich, 'Fichte's Original Insight', trans. David R. Lachtermann, in Darrel E. Christiansen *et al.*, *Contemporary German Philosophy*, vol. 1 (University Park: Pennsylvania State University Press, 1982).

22 Habermas, *Truth and Justification*, p. 201.

23 I am here alluding to John McDowell's proposition in *Mind and World* (Harvard: Harvard University Press, 1994), p. 67 that, for the absolute idealist, 'capacities of spontaneity' are 'in play all the way out to the ultimate grounds of empirical judgements'.

24 Habermas, *Truth and Justification*, p. 204.

25 Ibid., p. 203.

26 Frederick C. Beiser, *German Idealism: The Struggle Against Subjectivism, 1781–1801* (Cambridge, Mass.: Harvard University Press, 2002), pp. 1–2. Recent representatives of this view include Robert Solomon, *Continental Philosophy since 1750: The Rise and Fall of the Self* (Oxford: Oxford University Press, 1988) and Karl Ameriks, *Kant and the Fate of Autonomy* (Cambridge: Cambridge University Press, 2000).

27 Beiser, *German Idealism*, p. 5.

28 Ibid.

29 For my attempt to respond to Beiser's study, see Espen Hammer, 'The Legacy of German Idealism', *British Journal for the History of Philosophy* 11:4 (2003), pp. 521–35.

30 The theory of systemic intervention or 'colonization' of the life-world, which should be understood as Habermas's take on Hegel's positivity-thesis, is developed in great detail in the second volume of *The Theory of Communicative Action*, trans. Thomas McCarthy (London: Heinemann, 1984).

31 See Kenneth Baynes, *The Normative Grounds for Social Criticism* (Albany N.Y.: SUNY Press, 1992) for an excellent discussion of Habermas's Kantianism.

32 The development of formal pragmatics, or what he sometimes calls universal pragmatics, has a long and intricate history in Habermas's work. For the most important stations along the way, see 'Wahrheitstheorien' in *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns* (Frankfurt: Suhrkamp, 1984), pp. 127–83; *The Theory of Communicative Action*, trans. Thomas McCarthy (Boston: Beacon Press, 1987), vol. 1, pp. 273–337; and all the collected essays in *On the Pragmatics of Communication*, trans. Maeve Cooke (Cambridge, Mass.: MIT Press, 1998).

33 For the theory of discourse ethics, see the essays in Habermas, *Moral Consciousness and Communicative Action*, trans. Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, Mass.: MIT Press, 1990).

34 For representative statements of Apel's position, see 'Das Apriori der Kommunikationsgesellschaft und die Grundlagen der Ethik', in *Transformation der Philosophie* (Frankfurt: Suhrkamp, 1973), pp. 358–436; and "Sprechakttheorie und transzendentale Sprachpragmatik: Zur Frage ethischer Normen', in *Sprachpragmatik und Philosophie* (Frankfurt: Suhrkamp, 1976), pp. 10–173.

35 Habermas refers to 'a complementary relation' between philosophy and empirical theory. See his *Moral Consciousness and Communicative Action*, p. 39: 'The empirical theory presupposes the validity of the normative theory it uses. Yet the validity of the normative theory is cast into doubt if the philosophical reconstructions prove to be unusable in the context of application within the empirical theory'.

36 In order to demonstrate that the rules of discourse are 'inescapable presuppositions', Habermas introduces the notion of performative contradictions. A performative contradiction occurs when a particular speech act k (p) rests on noncontingent presuppositions whose propositional content contradicts the asserted proposition p. In one of Habermas's examples, the assertion 'Using lies, I finally convinced H that p' is said to be nonsensical and 'revisable' to 'Using lies, I finally talked H into believing that p.' (Habermas, *Moral Consciousness and Communicative Action*, p. 90). The point here is that convincing someone of something requires the offering of

justifications. If H lies, then he does not permit his proponent to form a rational conviction based on the consideration of reasons. He has talked H into something but not 'convinced' him. In my view, this simple analysis is not persuasive, and it hardly demonstrates that participants in argumentation *must* assume that rational persuasion cannot occur on the basis of lies. When Colin Powell presented evidence to the General Assembly of the United Nations to the effect that Iraq had weapons of mass destruction, he succeeded perfectly well in rationally convincing most of its members that this was true. I see no problem in saying that the evidence he presented (his 'good reasons') led to a reasoned agreement about the threat Iraq represented. Colin Powell and his associates knew, however, that what they did was precisely to 'use lies to convince H that *p*'. Without being able to demonstrate this claim here, I think similar problems beset all the rules of discourse. The more general lesson might be that speech acts are not governed in any strict sense by rules. Whether they make sense and achieve their illocutionary and perlocutionary aims depend ultimately on the concrete relation between speaker and hearer, and on the specific rhetorical and semantic context in which they are being made.

37 See the opening remarks on Hegel in Habermas, *Knowledge and Human Interests*, pp. 24–28.

38 Hegel, *Phenomenology of Spirit*, p. 66. Hegel's dialectic involves several different steps but here is a passage that captures the gist of the critique:

> They speak of the existence of *external* objects, which can be more precisely defined as *actual*, absolutely *singular*, *wholly personal*, *individual* things, each of them absolutely unlike anything else; this existence, they say, has absolute certainty and truth. They *mean* "this" bit of paper on which I am writing – or rather have written – "this"; but what they mean is not what they say. If they actually wanted to *say* "this" bit of paper which they mean, if they wanted to *say* it, then this is impossible, because the sensuous This that is meant *cannot be reached* by language, which belongs to consciousness, i.e. to that which is inherently universal. In the actual attempt to say it, it would therefore crumble away; those who started to describe it would not be able to complete the description, but would be compelled to leave it to others, who would themselves finally have to admit to speaking about something which *is not*.

39 Robert Pippin discusses some of them in 'You Can't Get There from Here', in Beiser (ed.), *The Cambridge Companion to Hegel*, pp. 52–85. For a defense of the priority of the *Science of Logic* in Hegel, including the *Phenomenology*, see Stanley Rosen, G. W. F. Hegel: An Introduction to the Science of Wisdom (New Haven: Yale University Press, 1974).

40 See Robert Pippin, *Hegel's Idealism: The Satisfactions of Self-Consciousness*, p. 108.

41 Habermas has a long story to tell about how the 'critical sciences' – Marxian critique of ideology and Freudian psychoanalysis in particular – by virtue of the very rationality they display, embody an interest in emancipation. He also, I think much more problematically, tries to show that the interest in emancipation is grounded in more deep-seated anthropological facts about the reproduction and self-formation of the human species.

42 Habermas, *Knowledge and Human Interests*, pp. 17–18.

43 Ibid., pp. 379–80.

> The studies I published in *Knowledge and Human Interests* suffer from the lack of a precise distinction … between reconstruction and "self-reflection" in a critical sense. It occurred to me only after completing the book that the tradi-

tional use of the term "reflection," which goes back to German Idealism, covers (and confuses) two things: on the one hand, it denotes reflection upon the conditions of the capacities of a knowing, speaking and acting subject as such; on the other hand, it denotes reflection upon unconsciously produced constraints to which a determinate subject (or a determinate group of subjects, or a determinate species subject) succumbs to in its process of self-formation. In Kant and his successors, the first type of reflection took the form of a search for the transcendental ground of possible theoretical knowledge (and moral conduct). . . . In the meantime, this mode of reflection has also taken the shape of a rational reconstruction of generative rules and cognitive schemata. Particularly the paradigm of language has led to a reframing of the transcendental model.

44 Theodor W. Adorno, *Negative Dialectics*, trans. E.B. Ashton (New York: Continuum, 1973), p. 23. 'Idealism – most explicitly Fichte – gives unconscious sway to the ideology that the not-I, *l'autrui*, and finally all that reminds us of nature is inferior, so the unity of the self-preserving thought may devour it without misgivings'.

45 Habermas makes this claim most explicitly in *The Theory of Communicative Action*, vol. 1, pp. 366–91, esp., p. 387: 'The critique of instrumental reason conceptualized as negative dialectics renounces its theoretical claim while operating with the means of theory'.

46 Adorno, *Negative Dialectics*, p. 5.

47 This, of course, is hardly an original claim. For a classical but essentially affirmative expression of the same point, see Ernst Cassirer, *Substanzbegriff und Funktionsbegriff: Untersuchungen über die Grundfragen der Erkenntniskritik* (Berlin: Bruno Cassirer, 1910).

48 It should be mentioned that Adorno does make gestures that point beyond the model of immanent critique. In 'Cultural Criticism and Society', in *Prisms*, trans. Samuel and Shierry Weber (Cambridge, Mass.: The MIT Press, 1997), p. 20, he claims that immanent critique 'remains imprisoned within the orbit of that against which it struggles'. There must be a certain freedom in regard to culture, for (p. 29) 'without consciousness transcending the immanence of culture, immanent criticism would be inconceivable'. Adorno remains, however, skeptical of the idea that critique can ever take place from a purely transcendent standpoint. It will always have to refer to the epistemic criteria that mark a specific object or object-domain. In a recently edited letter to his friend Gershom Scholem, Adorno (as quoted in Mauro Bozzetti, 'Hegel on Trial: Adorno's Critique of Philosophical Systems', in Nigel Gibson and Andrew Rubin (eds), *Adorno: A Critical Reader* (Oxford: Blackwell, 2002), p. 296) writes that 'I remain true to the *Phenomenology of Spirit* in my view that the movement of the concept, of the matter at hand, is simultaneously the explicitly thinking movement of the reflecting subject'.

49 Readers familiar with Adorno's thinking will hopefully recognize in this an appeal to the *Bildverbot* – the prohibition of graven images – which operates as a principle in this philosophy.