Jürgen Habermas

Article for *History of Continental Thought, Volume 6: Poststructuralism and Critical Theory: The Return of Master Thinkers,* ed. Alan D. Schrift (Acumen Press: forthcoming).

By Christopher F. Zurn, Department of Philosophy, University of Kentucky

This is a significantly longer version than will be published, and is still a draft. <u>Please do not cite without my permission!</u> I can be reached by email at: cfzurn [at] uky.edu.

Jürgen Habermas. German. **B:** June 18, 1929, Düsseldorf, Germany; **Educ:** Göttingen, Zürich, Bonn (1949-54); Promotion (~Ph.D.), Bonn (1954); Habilitation, Marburg (1961); **Infls:** *Kant, Hegel, Schelling, Marx, Nietzsche, Freud,* Peirce, Mead, Dilthey, Weber*, Durkheim, Horkheimer, Adorno;* **Appts:** Institut für Sozialforschung, Frankfurt (1956–59), U Heidelberg (1961–64), U Frankfurt (1964–71), Max-Plank Institut, Starnberg (1971–81), U Frankfurt (1975–94), Northwestern U (1994–2004)

Major Works:

Strukturwandel der Öffentlichkeit: Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft. Berlin: Luchterhand, 1962. English Translation: *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society.* Trans. Thomas Burger and Frederick Lawrence. Cambridge, MA: The MIT Press, 1989.

Erkenntnis und Interesse. Frankfurt: Suhrkamp, 1968. English Translation: *Knowledge and Human Interests.* Trans. Jeremy Shapiro. Boston: Beacon, 1971.

Legitimationsprobleme im Spätkapitalismus. Frankfurt: Suhrkamp, 1973. English Translation: *Legitimation Crisis.* Trans. Thomas McCarthy. Boston: Beacon, 1975.

Theorie des kommunikativen Handelns, Band 1: Handlungsrationalität und gesellschaftliche Rationalisierung and *Band 2: Zur Kritik der funktionalistischen Vernunft.* Frankfurt: Suhrkamp, 1981. English Translation: *The Theory of Communicative Action, Volume 1: Reason and the Rationalization of Society* and *Volume 2: A Critique of Functionalist Reason.* Trans. Thomas McCarthy. Boston: Beacon, 1984 and 1987.

Moralbewusstsein und kommunikatives Handeln. Frankfurt: Suhrkamp, 1983. English Translation: *Moral Consciousness and Communicative Action.* Trans. Christian Lenhardt and Shierry Weber Nicholsen. Cambridge, MA: The MIT Press, 1990.

Der philosophische Diskurs der Moderne: Zwölf Vorlesungen. Frankfurt: Suhrkamp,
       1985. English Translation: The Philosophical Discourse of Modernity: Twelve
       Lectures. *Trans. Frederick Lawrence. Cambridge, MA: The MIT Press, 1987.*
Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen
       Rechtsstaats. Frankfurt: Suhrkamp, 1992. English Translation: *Between Facts and*
       Norms: Contributions to a Discourse Theory of Law and Democracy. Trans.
       William Rehg. Cambridge, MA: The MIT Press, 1996.
Wahrheit und Rechtfertigung. Frankfurt: Suhrkamp, 1999. English Translation: *Truth and*
       Justification. Trans. Barbara Fultner. Cambridge, MA: The MIT Press, 2003.

    This essay seeks to give an overview of the development, central themes and
methods, and main claims of Jürgen Habermas's thought over the course of his career.
Given its incredibly wide thematic range, its pervasive influence in both public and
academic fora across diverse fields and disciplines, and the fact that it has taken many
different twists and turns (and reversals) over its course, any comprehensive
consideration of that body of thought will need to be selective. This essay selects through
three schematics. First, it periodizes Habermas's academic work into six phases:
presented-oriented philosophy of history, epistemology via philosophical anthropology,
the theory of communicative action, the discourse theory of morality, the discourse
theory of law and politics, and, systematic philosophical consolidation. Second, the essay
pays particular attention to the contexts of debate that have shaped Habermas's thought.
Finally, the essay attempts to trace three leitmotifs throughout Habermas's philosophical
career and corpus: a focus on *communication* as the immanent locus of the
transcendental, an insistence on the achievements of *reason* without ignoring the ravages
of modernity's one-sided employment of reason, and a conception of philosophy as
critical theory, that is, as reflective interdisciplinary theory oriented toward human
autonomy. The aim of the essay, then, is not so much to provide a systematic
presentation of Habermas's philosophy *simpliciter* as to provide an overview of some of
its main themes, problems, and claims by putting them into a biographical frame that
stresses the continuity of Habermas's central intuitions amidst his transformative
interactions with other thinkers and modes of thought.

I   **Present-Oriented Philosophy of History**

    As a 24 year old student of philosophy, Habermas had his first impact not with a
distinctive philosophical thesis or argument, but with a public intervention as a critic in
the sphere of letters. In 1953 he published a short newspaper piece criticizing
Heidegger's republication of his 1935 lectures, *Einführung in die Metaphysik,* which
were not only soaked through with rhetoric celebrating "the inner truth and greatness" of
National Socialism, but also attempted to align the question of Being itself with the
ascendancy of German fascism.[1] What shocked Habermas about these lectures was that

---

[1] Jürgen Habermas, "Mit Heidegger gegen Heidegger denken. Zur Veröffentlichung von
Vorlesungen aus dem Jahre 1935" *Frankfurter Allgemeine Zeitung,* July 25, 1953, pp.
67–75; English translation: "Martin Heidegger: On the Publication of the Lectures of
1935," trans. William S. Lewis, in *The Heidegger Controversy: A Critical Reader,* ed.
Richard Wolin (New York: Columbia University Press, 1991), pp. 190–97.

they were republished with no expression of regret or explanation, no acknowledgement of the painful truth of the horrors of the Third Reich, no admitting of political mistake or moral remorse. He treated this silence not simply as a mark against Heidegger, but as indicative of a general, and quite troubling, amnesiatic silence across post-war German culture, a constant evasion of "the problem of the prehistory of fascism."[2] The basic intellectual charge leveled in that short piece – that the underlying thought structure and content of Heidegger's philosophy did not undergo a "turn" from the earlier to the later work motivated by internal philosophical reasons but rather only rhetorical re-packaging in response to contemporary politics– remained constant throughout Habermas's published considerations of Heidegger's legacy across the decades.[3]

Beyond the specifics of this intellectual historiography, there are three signposts to future work already broached in this short piece. First, the article marks the end of Habermas's time as a through-going follower of Heidegger's thought,[4] and signals his firm commitment to the Enlightenment ideals of "individualistic egalitarianism" and anti-nationalistic "cosmopolitanism," coupled with a distinctive acknowledgment of the ambiguous character of reason, "the dialectical plasticity of modern development."[5] Second, it also marks the end of Habermas's belief that scholarly issues of pure philosophy could be neatly separated from concrete political issues of the present, and thereby marks the beginning of one of the major struggles throughout Habermas's career: how to engage both in scholarly academic work that makes a scientist's claim to universal validity and in practical political interventions that make claims merely as one citizen among others, without either having one activity simply collapse into the other or

---

[2] Habermas, "Martin Heidegger," p. 191.

[3] Beyond the 1953 article, see also two pieces published in 1959, "Die große Wirkung. Eine chronistische Anmerkung zu Martin Heideggers 70 Geburtstag" and "Ein anderer Mythos des Zwanzigsten Jahrhunderts," both collected in *Philosophisch-politische Profile* (Frankfurt: Suhrkamp, 1971), pp. 76–92, and "Work and Weltanschauung: The Heidegger Controversy from a German Perspective," trans. John McCumber in *The New Conservatism: Cultural Criticism and the Historians' Debate* (Cambridge, MA: The MIT Press, 1989), pp. 140–72. Further elaborations of Habermas's account of Heidegger's philosophy and its influence can be found throughout *The Philosophical Discourse of Modernity.*

[4] In "Martin Heidegger," Habermas claims that *Sein und Zeit* was "the most significant philosophical event since Hegel's *Phänomenologie*" (191) and closes with the admonition to "Think with Heidegger against Heidegger" (197). In interviews from the 1970s and 80s collected in *Autonomy and Solidarity: Interviews with Jürgen Habermas,* ed. Peter Dews (New York: Verso, 1992), he repeatedly refers to the importance of Heidegger to his early philosophical development: "Heidegger, in whose philosophy I had been living" (80); "The most powerful systematic impulse [in my philosophical studies] came from the early Heidegger" (147); "you may understand better what it meant for me, as someone who had been a thoroughgoing Heideggerian for three or four years, to read Marcuse for the first time" (189); "The two pre-eminent thinkers [in Germany after the war]—who determined the direction of my philosophical interests—were Heidegger and Gehlen" (192).

[5] Habermas, "Martin Heidegger," p. 196.

being unreflective about the inevitable cross-pollination of the two activities. Third, it marks the beginnings of Habermas's attempt to carry on philosophy not as a disconnected scholarly endeavor, but rather as a present-oriented philosophy of history, one taking seriously the left-Hegelian tradition of attempting a comprehensive assessment of the promise and pitfalls of the distinctive form of Western rationalization. Thus, as Habermas was finishing his 1954 dissertation on the problems in Schelling's account of the role of the absolute in history, a dissertation strongly influenced by Heidegger, he added a long "introduction setting late German Idealism in relation to Marx."[6]

　　　After the completion of his dissertation, Habermas worked for two years as a left-wing journalist writing on social issues before he became the personal assistant of Theodor Adorno at the *Institut für Sozialforschung* in Frankfurt. During the next few years, in addition to sociological work on the post-war German academic environment concerning issues such as students and their political attitudes, Habermas was also occupied with philosophically comprehending and assimilating what he had encountered first from a narrowly political point of view: the Marxist project of a critical theory of society, especially as it had been transformed and updated in the Western tradition of Hegelian Marxism starting with Georg Lukács and continuing in the work of the so-called "Frankfurt School" of critical theory by, among others, Adorno, Max Horkheimer and Herbert Marcuse. What interested Habermas in this tradition was neither the specifics of Marx's theory of capitalism nor his philosophy of history, but rather the prospect opened up in the early Marx of continuing the critique of modernity set in motion by German Idealism and Romanticism in the form of an account of a one-sided exploitation of the potentials of reason and rationalization. In quick succession appeared two works that would bring together the two already-exposed leitmotifs of a critical theory of society and an ambiguous attitude toward the promise and peril of modern reason, with the third motif of a focus on communication and communicative interaction as the immanent locus of context-transcending ideals.

　　　In his *Habilitationsschrift,* completed in 1961 under Walter Abendroth in Marburg, *The Structural Transformation of the Public Sphere*, Habermas pursued a relatively straightforward and delimited socio-historical study. The book focused on a central organizing category of liberal capitalist societies: the "public sphere" of humane letters and opinion where an interested public of private citizens comes together to exchange reasons, ideas, and arguments coalescing into a determinate public opinion. It traced how the public sphere first arose in the 18th century, anchored in new institutions such as widely-distributed newspapers, coffee houses, salons, and civil associations, was then institutionally changed by the rise of commercial journalism in the early 19th century, and was finally permanently restructured by the development of mass welfare-state democracies into a realm dominated by the mass media as platform for advertising to a culture-consuming public. It is, in other words, an historical investigation into the rise and degeneration of new forms of communicative interaction and their specific institutional preconditions. But the book is also a methodologically sophisticated interdisciplinary theory with emancipatory intent. It marshaled a welter of social-scientific evidence to explain how and why public communications amongst independent citizens originally played a central normative role in legitimating liberal democratic

---

[6] Habermas, *Autonomy and Solidarity,* p. 148.

states, only to be undermined and transformed into the manipulative manufacture of public opinion by social powers and special interests looking to direct the consumer behavior of politically disinterested subjects. In short, in revealing both the normative ideals embedded in the historical practice of the political public sphere, and the ways in which the celebration of those ideals became ever more ideological and false as the public sphere itself changed, Habermas's book showed that questions of political philosophy concerning the legitimacy of liberal democracy must be systematically connected to questions concerning the specific socio-historic contexts and functions of the institutions and social arrangements in which those ideals are embedded. *Structural Transformation* thus definitively introduced the three themes of communication, reason, and critical theory that would form the backbone of the next five decades of Habermas's work, even as many of the surrounding philosophical and socio-theoretical details would undergo significant transformation.

   The next year, 1963, saw the publication of a collection of essays entitled *Theorie und Praxis. Sozialphilosophische Studien* continuing in the same vein, but approaching its subject by reworking the themes of classical political theory – especially those of social contractarianism, natural law liberalism, constitutional republicanism– from within the framework of an updated but still recognizably Marxist philosophy of history.[7] In particular, it sought to explore the changing physiognomy of political and social thought from its development in classical Greece through its transformation first under the modern conditions at the time of the liberal capitalist revolutions, and then again under the changed historical conditions of mass, welfare-state democracies organized by rational bureaucracies oriented toward technical progress.

II   Epistemology via Philosophical Anthropology

   Habermas received his first professorship at Heidelberg in 1961, thanks in large part to the efforts of two prominent students of Heidegger to hire him there: Karl Löwith and Hans-Georg Gadamer. Even more important for his development than this personal connection, however, was the 1961 publication of two books that decisively influenced all his future work by re-orienting his considerations of everyday, ordinary human communication from being one among several interesting topics to being the absolute center of his philosophical thought – a position, even through many changes, which it has retained to this day. As he himself put it, Gadamer's "*Wahrheit und Methode,* together with [Ludwig Wittgenstein's] *Philosophischen Untersuchungen* which appeared at the same time, gave the stimulus to the thoughts which one could fully describe as the

---

[7] Jürgen Habermas, *Theorie und Praxis. Sozialphilosophische Studien* (Neuwied-Berlin: Herman, 1963). Partially translated into English, with essays added from 1966, 1967 and 1971 as *Theory and Practice,* trans. John Viertel (Boston: Beacon Press, 1973). For further bibliographic details concerning the contents of various collections in both German and English, see the excellent bibliography through 1980 compiled by René Görtzen and Frederik van Gelder in Thomas McCarthy, *The Critical Theory of Jürgen Habermas* (Cambridge, MA: The MIT Press, 1978), pp. 442–64, and, the comprehensive bibliography through 1995 edited by Demetrios Douramanis, *Mapping Habermas from German to English: A Bibliography of Primary Literature 1952–1995* (Sydney: Eurotext, 1995).

'linguistic turn of critical social theory.'"[8] At the same time, his friend and frequent collaborator Karl-Otto Apel introduced him to American pragmatist thought, especially the work of Charles Sanders Peirce, John Dewey, and George Herbert Mead. He has remarked that "from the outset I viewed American pragmatism as the third productive reply to Hegel, after Marx and Kierkegaard, as the radical-democratic branch of Young Hegelianism, so to speak."[9]

The ten years from his Heidelberg appointment, through his double professorship in Frankfurt in philosophy and sociology (taking the place of Max Horkheimer) in 1964, to his resignation from that post in 1971, was incredibly prolific and saw the development of a fully articulated, comprehensive research program of social and political philosophy. In retrospect what is remarkable is that most of the major topoi of Habermas's philosophical career – the critique and diagnosis of modernization processes, the aim to grasp the place and import of science and technology in our lifeworld, the methodological clarification of critical social theory, the endeavor to update the substantive claims of critical social theory under changed historical conditions, the differences between market and bureaucratic versus communicative modes of sociation, the import of a pragmatic consideration of language in its everyday use, the diversity of forms of reason and its claims to universal validity, and others – were already broached during this period, even as most of the specific content of his substantive claims, arguments, and theories concerning those topoi would undergo significant if not radical transformation in the next period. Given constraints, the treatment here of this period is especially selective, focusing only on Habermas's general attempt to develop a multi-facetted concept of reason, one that resists both reductivist attempts to understand all employments of reason under the aegis of a single one-sided model, and complementary skeptical whole-sale rejections of a specific use of reason in the name of what its use has occluded, distorted, or oppressed.

In the early 1960s, spurred by an unproductive conference encounter between Adorno and Karl Popper concerning the proper understanding of science and of social scientific research in particular, Habermas wrote a series of important articles attempting to rebut scientistic and positivistic understandings of scientific research.[10] Two of the central issues that he raised concerned scientism's claim to the exclusive rationality of

---

[8] Jürgen Habermas, "Hommage an Hans-Georg Gadamer. Er erforschte '*Wahrheit und Methode*' der philosophischen Erkenntnis—Am 11. Februar wird er 100," in *Der Tagessspiegel,* January 2, 2000, quoted in the short biography by Rolf Wiggershaus, *Jürgen Habermas* (Hamburg: Rowohlt Verlag, 2004), p. 59 (my translation).

[9] Habermas, *Autonomy and Solidarity,* p. 148.

[10] Jürgen Habermas, "Dogmatismus, Vernunft und Entscheidung. Zu Theorie und Praxis in der verwissenschaftlichen Zivilisation," in *Theorie und Praxis,* pp. 307–36; "Analytische Wissenschaftstheorie und Dialektik. Ein Nachtrag zur Kontroverse zwischen Popper und Adorno," in *Zeugnisse. Theodor W. Adorno zum sechzigsten Geburtstag,* ed. Max Horkheimer (Suhrkamp: Frankfurt, 1963), pp. 473–503; "Gegen einen positivistisch halbierten Rationalismus. Erwiderung eines Pamphlets," in *Kölner Zeitschrift für Soziologie und Sozialpsychologie* v. 16, #4 (1964): 636–59; *Zur Logik der Sozialwissenschaften,* in *Philosophische Rundschau* v. 14, supplement 5 (1966–67); "Technik und Wissenschaft als Ideologie," in *Merkur* v. 22, #234 (1968), pp. 591–610.

scientific knowledge in the hypothetical-deductive form, and, positivism's claim to the value neutrality of both science and the philosophy of science. On the first point, Habermas employed several different strategies to argue against the notion that the combination of logical-analytic truths with either naïve empiricist references to a theory-independent world of states of affairs or with more sophisticated critical rationalist references to objective facts exhaust the domain of claims that could be critically evaluated. In essence, Habermas argued that the domain of cognitive claims went well beyond the sphere delimited by a narrowly scientistic self-understanding of scientific research. One of the most powerful arguments here is Habermas's claim that the standards and norms of scientific practice extolled by positivists – and even their celebration of rationality itself – cannot be justified by reference to only those truth-makers their own meta-theory commits them to recognizing (i.e., empirical states of affairs or objective facts). The specific meta-theoretic endorsement of certain scientific norms of practice above others and the basic preference for science itself can only be justified in the light of specific human interests or values, yet scientism is committed to the claim that all evaluative concerns about human interests and values are themselves, at bottom, incapable of rational justification. Thus Habermas attempted to show that the scientistic exclusion of normative issues from the domain of rational assessment undermines the very normative standards that scientism presupposes. Concerning the second point of positivism's supposed value neutrality, Habermas's move was structurally similar: not only is such value neutrality false (since positivism has a preference for scientific rationality that positivism itself cannot justify), but it is also ideological. With this latter claim, Habermas connected back up to the tradition of critical theory with the idea that the exclusivist preference for scientific rationality in effect pre-loads the dice in favor of either decisionism or technocracy in various socio-political domains. For either the policy pronouncements of politicians, bureaucrats and administrators are legitimated on the basis of their expert knowledge which is epistemically inaccessible to a democratically debating public sphere, or their policy choices are legitimated only by the plebiscitary power conferred on them in the light of their charismatic capacity to convince a supine electorate of their decisive personal will and trustworthy gut instincts. In either case, the liberal and republican ideals of political opinion and will formation depending on a public and democratic process of debate and discussion leading to the best or most rational outcome available – the cornerstone Enlightenment ideal of the democratic public sphere – are ideologically made to seem mere quaint ideas of a pre-technological world.

It is important to emphasize that the various theoretical critiques Habermas leveled at scientism and positivism were always directed at false theories and understandings of science, not at the actual practices and results of scientific research which would require detailed, individualized assessment. For forgetting this distinction might lead one to an equally deluding but opposite attitude toward science and technology from that of positivism: the idea that most all of the ills of contemporary society can be laid at the feet of the dominance of scientific and technological ways of thinking simpliciter. Whether in the form of Heidegger's eschatology of Being whereby we have fallen into an almost hopelessly instrumentalized form of existence, or in Marcuse's Heideggerian Marxism that combined that history of metaphysics with the claim that all forms of contemporary social and political domination can be traced not to

capitalism but to means-ends rationality itself, variations of all-encompassing technocracy theories went beyond noting the limitations of the empirical-analytic sciences to wholesale rejections of modern science itself as well as its techno logical byproducts. Agreeing with the critics of technical thinking that it alone cannot claim the exclusive mantle of reason, Habermas nevertheless insisted – against, for instance, Marcuse's romantic calls for replacing scientific and technological progress with a new, "humane" science and technology – on acknowledging the unsurpassable achievements of modern exact research. Such achievements are not a mere historical accident, nor can they be disposed with, at least as long as the human species has an interest in its own self-preservation and in its increasing independence from material need. The question then is how to understand and deal with the one-sided institutionalization of technical thought to the exclusion of other forms of rational thought and communication.

One last critical encounter during this period deserves mention: Habermas's *Auseinandersetzung* with hermeneutics, especially as powerfully formulated by Gadamer. As mentioned, Habermas was an early and important defender of hermeneutic methods in the social sciences, and he agrees with many of the foundational ideas of Gadamer's account of hermeneutics: the stress on everyday ordinary language as the medium of social life, the importance of background prejudgments and prejudices (*Vorurteilen*) to any act of understanding, the social reproduction carried by inculcation into and increasing understanding of one's own cultural tradition, the historical specificity of traditions combined with their constitutive openness to learning from other traditions through the expansion of one's horizons, the commitment to dialogue as an integral part of practical reasoning and ethical life, the ineradicably practical orientation of social investigation, the methodological impossibility of a full grasp of symbolically-structured social domains through quantitative, objectivating empirical-analytic social science, the essentially historical character of sociological understanding against the positivist illusory hope for general ahistorical covering laws of explanation, and so on.

What then separates the two? It is not only a basic difference of temperament: the more conservative Gadamer comfortable with the truths of tradition versus the more radical Habermas suspicious of accepting anything on the mere authority that it has been long accepted, the theorist of judgment versus the theorist of reflection, the contextualist versus the universalist, the humanist versus the enlightener. The most important issue concerns precisely the status and scope of philosophical hermeneutics' claim to universality. While Gadamer insists that no form of experience, no form of science or knowledge can be excepted from the methodological constraints of hermeneutics since the community of language and tradition simply is the medium of the human form of life, Habermas holds out for the possibility of modes of analysis that reveal systematic forms of constraint or distortion operating, as it were, behind the backs of ordinary-language using social participants. Thus while Gadamer insists on absolutizing the form of understanding theoretically articulated by hermeneutics, Habermas insists that there are other employments of reason which can yield important insights. In particular, Habermas is concerned with maintaining methodological space for social scientific research which allows one to grasp cultural traditions and ordinary languages as one among a host of interacting determinants crucial to grasping social reality. For if it is possible that language and tradition often themselves embed relations of domination and force, then one would need access to a type of ideology critique that could show how tradition itself

has been partially constituted by material relations of power. Furthermore, hermeneutics absolutized would deny the insights of functionalist sociology, which points to economic and political processes that operate according to the quantifiable logics of profit and power, rather than the qualitative logics of culture and meaning. And the same can be said for psychoanalysis, which promises insights into the way in which idiosyncratic semantic contents are externally caused by internalized repression and allied dynamic mechanisms. Even apart from such worries about domination and emancipation, theories such as Chomsky's concerning universal grammar and Piaget's concerning universal stage-sequential process of cognitive development have the potential to transcend the interpretive limitations of specific languages and traditions. Finally, Habermas does not want to renounce the possibility of developing a philosophy of history that goes beyond the history of ideas to also acknowledging the importance of changes in the structures of material reproduction through human labor and in the organization of political power and domination, a renunciation apparently called for in the name of a hermeneutic idealism that would insist on seeing all social phenomena in culturalist terms all the way down. Such examples of empirical social science that attempt to theorize causal mechanisms and generalize their results across various traditions can then not only provide important insights into social reality; they also promise, to the extent they are veridical, access to truths that are not contextually-bound in the same way as hermeneutic insights gained from the internal perspective of a way of life.

These various critiques of scientism and positivism, technocracy theories, and philosophical hermeneutics all share the feature of pointing out a one-sided absolutization of important insights and insisting on the plurality of the uses and methods of reason. But as an inveterate systematic thinker, Habermas was never committed to a sort of piecemeal eclecticism or bricolage. Rather, he sought the positive articulation of a comprehensive theory that would integrate the various legitimate forms of knowledge into a comprehensive critical social theory. His 1965 inaugural address on assuming Horkheimer's chair in philosophy and sociology in Frankfurt announced his intention to provide an *epistemological* foundation for an integrated, interdisciplinary theory with emancipatory intent, an intention that was brought to fruition in the masterful 1968 *Erkenntnis und Interesse*. Critically evaluating the epistemological programs of a diverse range of philosophers including Kant, Fichte, Hegel, Marx, Comte, Mach, Nietzsche, Pierce, Dilthey, and Freud, Habermas attempted to show how their insights and limitations could best be understood from a basic anthropological perspective. Rejecting the classical epistemological doctrine that pursuing practical interests is antithetical to achieving knowledge, Habermas maintained that all forms of epistemic inquiry should be seen as motivated by one of three anthropologically-basic, fundamental human interests: the *technical* interest in the prediction and control of the natural environment, the *practical* interest in the reproduction of the social form of life achieved through intersubjective communication, and the *emancipatory* interest in freeing our selves and our societies from all forms of falsely naturalized but changeable constraints. The audacious claim of the book was that these three interests operated as constitutive conditions of possibility in the organization of three different forms of inquiry – empirical-analytic sciences, historical-hermeneutic sciences, and critical sciences – where each form of inquiry is internally structured – in its basic categories, in its grasp of its own object domain, in its logic, and in its form of characteristic validity claim – by its

distinct underlying fundamental human interest. Further, each interest constitutively structures a central element of human social life: work, language, and power respectively. Thus the affinity of modern science with technological achievements is no accident, for when the natural and social sciences aim at general empirical knowledge organized into theoretical hypothesizes making truth claims, they are categorically structured by the human interest in our interchange with the natural world, an interchange regulated by the relevant forms of social labor. The close connections between historiography, the humanities, and interpretive social sciences and the classical pursuits of morality, ethics, rhetoric and politics are likewise no accident, for the former aim at the theoretically-articulated understanding and self-understanding of practically-oriented human agents as they attempt to decide on the best course of action in the light of their socialization into, and reproduction of, linguistically and normatively structured domains of social life. Finally, according to Habermas, the otherwise surprising structural similarities between psychoanalysis, ideology critique, and critically reflective philosophy are best understood by seeing that all three are forms of inquiry shaped by the interest in emancipation from falsely naturalized, but actually changeable, power relations not otherwise evident or obvious on the surface of psychological and social life.

Given the excited critical reception of *Knowledge and Human Interest,* it may appear surprising that Habermas never carried through the basic research program. While he continued to express confidence in the basic outlines of the argument, he acknowledged that he would write it differently were he to do it again. Without getting too deeply into details, by 1973 it was clear that Habermas had significant reservations about the book and had attempted to resolve them – not by revisiting the project, but by developing a new research program tailored to attend to the problems along the way.[11] One of the most significant problems concerned the third form of epistemic inquiry: the status and aims of critical social theory itself. Habermas later acknowledged that, in the book's attempts to revive the insights of the German idealist tradition of reflective self-critique developed from Kant to the early Marx, it employed a systematic ambiguity in the use of the concept of reflection between the Kantian idea of reason's reflection upon its own necessary conditions of possibility and the young Hegelian idea of persons' and societies' reflection upon otherwise inconspicuous forms of domination and power. While the first form of reflection aims at grasping the universal generative structures and rules of a particular use of reason, the second form aims at emancipation from systematically constraining, but unacknowledged forces and powers, whether intrapsychic, ideological, social or material. But how can the same activity – critical social theory – aim both at liming the necessary conditions of human inquiry and at uncovering the socio-historically contingent features of modern social life that impede the realization of freedom, at the same time and with the same tools? Habermas needed to develop a much clearer picture of the various components of a critical social theory, how they related to one another, and the status of their respective claims to validity. Furthermore in providing, as it were, an epistemological prolegomena to critical theory, one that spelled out the philosophical possibility of a critical social theory, *Knowledge and Human Interests* did not offer such a theory: essentially a meta-philosophical treatise,

---

[11] Jürgen Habermas, "A Postscript to *Knowledge and Human Interests,*" in *Philosophy of the Social Sciences,* v. 3, #2 (1973): 157–89.

it did little to advance substantive socio-theoretic claims or diagnose the peculiar character of the present state of Western society and its relevant past with an emancipatory intent.

III   The Theory of Communicative Action

In 1971, Habermas accepted an academic researcher's dream job. The physicist and peace activist Carl Friedrich von Weizsäcker invited him to a directorship at the Max Plank Institute for Research into the Living Conditions of the Scientific-Technical World in Sternberg outside of Munich, enabling Habermas, together with at least 15 co-workers, the opportunity to reconstitute his research program on a new foundation, one thoroughly grounded in the latest results of diverse social-scientific domains and results. At least since the publication of *Knowledge and Human Interests,* Habermas had begun to be suspicious of the heavy argumentative burdens his program assumed in incorporating strongly Hegelian and metaphysical conceptions of notions such as truth, totality, and philosophy, and sought ways to make critical social theory as he understood it much more receptive to empirical research and methodologically open to empirical fallibility, though for some time he apparently thought the research program was still best framed in epistemological terms. Eventually, he became convinced that epistemology and methodological reflection were not the royal roads for critical theory, but that one should rather develop a substantive theory of society that would be able to show how communicative action is itself the immanent, practical locus of context-transcending reason, and that from an analysis of communicative action itself one could locate and justify the impetus toward an emancipated society, and thereby supply a new basis for a substantive critical theory.

Continuing to exercise his apparently limitless capacities for assimilating, comprehending, and systematizing entire research programs across all fields of social-scientific and humanistic investigation – witnessed earlier in his productive interactions and debates with the varieties of Western Marxism, modern political philosophy, analytic philosophy of science, German idealism, various forms of phenomenology, Gadamer's hermeneutics, American pragmatism, the varieties of psychoanalytic theory – Habermas in the late 60s and accelerating into the early 1970s was busy coming to terms with a multiplicity of cutting edge research: ethnomethodology and social phenomenology (especially as developed by Aaron Cicourel and Alfred Schutz[12]); Chomsky's theory of a universal, generative grammar; John Austin's pragmatic emphasis on speech acts and John Searle's systematization of speech act theory; the history of sociology (especially Weber); contemporary structural functionalist sociology and social-psychology (especially of Talcott Parsons and his followers); developmental psychology concerning cognitive competence (Jean Piaget) and moral reasoning development (Lawrence Kohlberg), and so on.

The next decade saw a remarkable proliferation of work – including the 1973 book *Legitimationsprobleme im Spätkapitalismus,* articles on cognitive and psychoanalytic psychology, on moral development, on ego identity, on social

---

[12] Cf. Aaron V. Cicourel, *Method and Measurement in Sociology,* New York: Free Press , 1964 and Alfred Schutz, *Der sinnhafte Aufbau der sozialen Welt,* Vienna: ??, 1932, English Translation: *The Phenomenology of the Social World,* Evanston, IL: Northwestern University Press, 1967.

psychology, on evolutionary theories of history, on the reconstruction of historical materialism, on communicative competence, on systematically distorted communication, on linguistic pragmatics and on interactive pragmatics, on truth, and many on individual philosophers and social theorists[13] – all capped by the appearance of Habermas's new magnum opus in 1981: the two-volume *Theory of Communicative Action.* Rather than work through all of this material historically, I will give an overview of the themes and central claims of the now mature critical social theory developed in this decade, organizing it under three themes: the linguistic turn in critical theory, the integration of systems theory and attendant diagnoses of the present, and the debate with poststructuralists and postmodernists over the meaning of modernity.

The Linguistic Turn

The most important component of Habermas's new version of critical theory – and the most recognizable one in its reception – is surely his focus upon language, specifically upon the basic structures evident in the *use* of language for purposes of intersubjective communication aimed at coordinating action. Taking off from studies in analytic philosophy of language that focus on language in its actual use (speech act theory), rather than upon semantic issues of fixing meaning or syntactic issues concerning the structures of sentences, Habermas aims to reconstruct the implicit yet highly developed know-how that competent linguistic communicators presuppose and rely upon when they engage in communicative action. He aims, then, at developing a formal pragmatics of language use, a theory that articulates the pre-theoretical knowledge, competences, and concepts employed by ordinary persons any time they endeavor to communicate with another person about something in the world in order to coordinate their individual actions.

A starting place for understanding the theory might be the distinction between two different ways in which one can employ language in order to achieve some intersubjective result, a distinction that ensues in two different fundamental social action types, that is, two different ways in which the disparate action plans of individuals can be coordinated. On the one hand, one might use language simply to convey certain information to another actor about the influence (direct or indirect) one has over the other, say in terms of imparting a threat or a promised reward to the other should s/he do as one wishes. In this case, Habermas claims, one is using language *strategically* and if the use is successful then the two agents will have coordinated their interaction according to the *strategic action* type. Typical examples include here not only manifest expressions of coercive power, but also various forms of mutual bargaining with reciprocal threat potentials, conscious uses of deception and manipulation, and even unconscious forms of deception due to what Habermas calls systematically distorted communication (e.g., socially-pervasive ideology or intrapsychic repression). On the other hand, one might use language to come to a mutual understanding with another person about something, for instance about a factual state of affairs in the objective world, or the prevailing norms in the social world, or the individuals' experiences in their respective subjective worlds.

---

[13] An important collection of essays from this period is Jürgen Habermas, *Zur Rekonstruktion des Historischen Materialismus,* Frankfurt: Suhrkamp, 1976, partially translated into English as *Communication and the Evolution of Society,* Trans. Thomas McCarthy, Boston: Beacon Press, 1979.

One absolutely essential aspect of such *communicative* uses of language is that their success hinges upon the ability of a respondent to take up a "yes" or "no" position on another's speech act offer, such that we can speak of *communicative action* when the coordination of persons' individual action plans is achieved through mutual agreement between them. Typical examples here can be seen when some ordinary course of action is disturbed, but can be repaired through mutual agreement on some state of affairs (e.g., the time of the meeting), the proper application of a relevant norm (e.g., the manager does have the authority to determine meeting times), or the character of an individual's experience (e.g., the misinterpretation of the manager's statement was a sincere mistake not a deceptive attempt to avoid the meeting). Although Habermas has repeatedly revised and reworked his formal pragmatics since its initial development in the early 1970s,[14] one crucial thesis has remained constant: the communicative use of language is fundamental, whereas other uses of language – strategic, fictional, figurative – are parasitic on or derivative from the properties and structures of communicative action. This means that, from the point of view of social theory, one must be able to understand the formal structures of communicative action in order to understand other forms of social integration – there is no possibility of radically bracketing ordinary language communication and the participant's perspective it requires social theory to adopt.

Communicative action is fundamentally intersubjective in the sense that it can only work to the extent that each individual agrees with the speech act offer made. Said otherwise, communicative action fails whenever one of the interlocutors rejects a claim made by another: each is assumed to be a competent actor who can assess the inherent *validity claims* made by others, and action coordination is only achieved when all involved come to a mutual agreement accepting the speech act offer. Furthermore, when individuals make validity claims in speech acts, whether implicitly or explicitly, they are thereby claiming that the content of their speech act deserves intersubjective recognition from others. Here Habermas distinguishes between four types of implicit validity claims made in each and every speech act: that the utterance is *comprehensible* or *intelligible* (semantically and grammatically well-formed), that the utterance is *true*, that the norms of social action invoked are *right*, and that the speaker is *truthful* or *sincere* in making the utterance. While the claim to comprehensibility is limited to the formation of the particular speech act, the other three types may be described as *universal* validity claims insofar as they involve an in-principle appeal to the notion that any competent agent would have to agree with the content of the claim, under suitable conditions for the evaluation and redemption of that kind of validity claim. Thus any time a speaker makes a communicative utterance, the speaker concomitantly makes four types of validity claims – even if, as is usual, only implicitly – and the hearer of the speech act may challenge the speaker on any of the four registers. Conversely, on accepting the speech

---

[14] Many of the early preparatory papers and lectures are collected in Jürgen Habermas, *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns* (Frankfurt: Suhrkamp, 1984). Two English language collections contain much of this work, as well as further revisions from the 1980s and 1990s: Jürgen Habermas, *On the Pragmatics of Social Interaction: Preliminary Studies in the Theory of Communicative Action,* trans. Barbara Fultner (Cambridge, MA: The MIT Press, 2001) and *On the Pragmatics of Communication,* ed. Maeve Cooke (Cambridge, MA: The MIT Press, 1998).

act without challenge, the hearer also commits him or herself to the various validity claims made in the utterance in the sense that he or she must also then be prepared to defend them against challenges. According to Habermas, it is precisely this intrinsic link between ordinary language use and the validity claims actors implicitly raise and accept that accounts for the illocutionary force of speech acts offers, or what he often calls "the binding/bonding force" of language. Communicative action works to coordinate action because in accepting the speech act offer of another, one also implicitly agrees that the universal validity claims raised by the other deserve intersubjective recognition. Individuals who come to a mutual understanding on a speech act are rationally motivated to carry through on their action commitments since their own agreement to the content of the utterance is freely made on the basis of their own individual insight into the propositional truth, normative rightness, and subjective sincerity of its content.

 Habermas's claim that social order is produced and reproduced through the consensus formation witnessed in communicative action might seem highly improbable. After all, not only is such a consensus ever threatened by new problem situations, new experiences, the different perspectives of different individuals, changing states of the natural and social worlds, and so on, but it is also fully contingent on the unforced agreement of social participants who can at any time refuse to say "yes" to a speech act offer. Here Habermas agrees with a host of 20th century theories – especially social phenomenology, hermeneutics, ethnomethodology, and ordinary language philosophy – that insist on the need for a massive background consensus to stabilize reciprocal understanding, and he adopts Husserl's concept of the lifeworld to explain how this unthematized background knowledge contains the shared meanings, beliefs, norms, and personality structures that absorb, as it were, the contingency built into communicative action. The lifeworld operates as a font of epistemic and practical certainties for interlocutors who can largely presuppose that interlocutors live "in the same world," that they do. Of course, when communicative interaction breaks down, it is possible for interlocutors to bracket ordinary interactions, explicitly focus on one, specific contested part of the lifeworld background, and engage in a distinctive kind of reflective argumentation that Habermas labels "*discourse*." Here interlocutors suspend their ordinary purposive orientations in a collective, more or less disinterested search for the truth of the matter – or for the normative rightness of the standards invoked, or for the degree of sincerity of the speaker – and they engage in a more demanding process of reason-giving under the supposition that consensus can only be achieved according to the "unforced force of the better argument."

 In distinction from earlier philosophical traditions that treat the concept of the lifeworld largely from an epistemological perspective, however, Habermas wants to employ the concept for social theory and so significantly broadens its application. In particular, by following the distinction between the three universal validity claims implicit in speech acts, Habermas proposes that there are three components of the lifeworld – cultural paradigms, legitimate orders, and personality structures – with each focused around a characteristic speech act type centrally thematizing one form of validity claim: constatives thematize truth claims, regulatives thematize normative rightness claims, and expressives thematize subjective sincerity claims. Furthermore, in modern complex societies, discourse itself has become reflective and taken on methodical institutional form in differentiated knowledge systems corresponding to the three

universal validity claims: science and philosophy systematically investigate propositions according to the logic of truth claims, law and morality systematically investigate illocutionary content according to the logic of rightness claims, and art, literature and criticism of taste investigate intentional and expressive content according to the logic of authenticity and sincerity claims.

Beyond its application to the sociological distinction between culture, society and personality, the theory of communicative action plays a central philosophical role in Habermas's oeuvre as the basis for his attempt to redeem rationality against various forms of skepticism. The cornerstone of this project is the notion of formal pragmatics. For while one might be interested in looking at the context-specific details of different societal lifeworlds, Habermas is interested in the deep, formal and invariant structures of all lifeworlds. Where philosophers traditionally sought to identify and justify the ideals of reason through speculative metaphysics, Habermas seeks to locate these ideals immanently in the very practices of communicative intersubjectivity by articulating the *idealizing pragmatic presuppositions* inevitably made by competent social actors when they engage in linguistic interaction. There are at least five centrally important idealizing presuppositions assumed by communicating agents: that individuals share a common objective world; that sameness of reference is maintained across interlocutors' speech acts; that interacting agents are rational, responsible, competent, and accountable; that validity claims have the sense of unconditionality or context-transcendence; and, that inclusive and open rational discourse alone is the unavoidable route to warranted justification. Furthermore once inquiry gets to the level of discourse, even more demanding and idealizing pragmatic presuppositions come in to play in structuring argumentative interactions aiming at the rigorous testing of validity claims: that no competent persons have been excluded from the conversation, that each has an equal voice, that each is free to speak without deception or self-deception, that no coercive influences are built into the processes of argument, that there are no operative time or decision pressures foreclosing necessary investigations, that only the unforced force of the better argument motivates rational acceptance of a claim, and so on. To be sure, all of these presuppositions are counterfactual in the sense that they are never fully realized in any concrete interaction, but they are nevertheless factually effective in structuring actual interactions. They can, in fact, be used normatively to critique any actually-achieved agreement as deficient from the point of view of the very standards of reasonability built into the practice itself. The pragmatic presuppositions of communication and discourse function thereby as immanent standards of a self-correcting learning process.

One prominent application of formal pragmatics has been to the traditional topic of epistemology: the nature of truth. Though I can't go into details here, it is worth noting that Habermas's account of truth has changed over time, such that we can point to at least three main phases. In all phases Habermas has rejected both correspondence theories of truth – for naively supposing linguistically and conceptually unmediated access to brute facts – and coherence theories of truth – for over-inflating the significance of linguistic mediation to the degree that they exclude the way in which truth claims are about states of affairs rather than merely about other statements. Habermas's first serious go at a theory of truth was in the 1973 paper "Wahrheitstheorien" where he laid claim to a "consensus" theory of truth: statements are true when they have been agreed to by all

dialogue participants under ideally extended conditions of justification.[15] The relation of this criterion of truth to Peirce's concept of an unlimited discourse community should be evident. In the second phase under the influence of Hilary Putnam's and Crispin Wright's work, Habermas sought to tighten up the earlier conception by invoking a notion of truth as idealized warranted assertability. Since 1999, Habermas has abandoned the purely epistemic theory of truth in favor of a version of what he calls pragmatic realism. The idea here is to attend to the different ways in which claims to truth function in everyday life and in reflexive discourses. In short, in our everyday dealings with the social and objective worlds we are firm realists, convinced of the unconditionality and context-transcending validity of truth claims. As Habermas puts it, "we do not walk onto any bridge whose stability we doubt."[16] Yet when we engage in reflective discourse about particular truth claims, for instance in scientific investigation, we realize that truth claims are epistemically tied to unavoidably linguistic practices of justification such that we take on the idealizing pragmatic presuppositions of discourse including the notions of ideal assertability to an unlimited communication community and of the inherent fallibility of all provisionally justified truth claims.

One can see in the general project for a formal pragmatics a flowering of what was previously an under-theorized concept in *Knowledge and Human Interests*: namely, the quasi-transcendental. For while formal pragmatics is aimed at the conditions of possibility of using language to coordinate actions with others – and is in this sense a continuation of Kantian transcendental philosophy aimed at unavoidable, universal features of language – its claims are distinctly rooted in an empirical analysis of actual language use by ordinary speakers – and is in this sense an *a posteriori* endeavor fallibilistically subject like all empirical knowledge to evidentiary testing. In this, Habermas sees his formal pragmatics as on a par with other *reconstructive sciences* such as Chomsky's universal grammar, Piaget's cognitive developmentalism, or Kohlberg's moral developmentalism. In each case, a reconstructive science aims to elucidate in theoretical terms the pre-theoretic know-how of competently interacting human agents, making empirically fallible claims to the universality of the underlying generative rules and structures of agentic competence. "After the pragmatic deflation of the Kantian approach, 'transcendental analysis' means the search for presumably universal, but only de facto inescapable conditions that must be met for certain fundamental practices or achievements."[17] If then Habermas's formal pragmatic claims are correct, he has located in linguistic intersubjectivity itself the immanent locus of the transcendent; he has fulfilled one of the critical desiderata for left-Hegelian thought that had eluded past versions of critical theory still caught up in the ceaseless back and forth between objectivism and subjectivism that was an artifact of a monological philosophical framework. In fact, we might say that critical theory according to Habermas's

---

[15] Jürgen Habermas, "Wahrheitstheorien," in *Wirklichkeit und Reflexion. Festschrift für W. Schulz,* ed. H. Fahrenbach (Pfüllingen: Press??, 1973), pp. 211–65. That Habermas was never happy with this working paper is evinced by the fact that he never let it be published in English translation.

[16] Jürgen Habermas, *Truth and Justification,* ed. and trans. Barbara Fultner (Cambridge, MA: The MIT Press, 2003), p. 39.

[17] Habermas, *Truth and Justification,* p. 86.

reformulation is composed of a grand synthesis of reconstructive social sciences aiming at universal, formal structures of any human form of life mediated by linguistic communication, along with other context-specific results of social scientific research as are necessary for the assessment of the specific form of life encountered in contemporary, developed Western societies.

Habermas's interest in reconstructive sciences such as Piaget's and Kohlberg's is not however only methodological, but includes the substance of their theories. In particular he contends that their notion of a stage-sequential, directional, progressive developmental logic can be employed not only in understanding the maturational processes of human individuals, but also be put to use in comprehending social evolution. Consider first the rich and variegated account of ontogenesis Habermas began developing in the early 1970s and continues to refine, that aims to integrate a number of different research paradigms used to study ontogenesis, particularly the cognitive developmental (especially Piaget and Kohlberg), psychosexual (especially psychoanalytic), and symbolic interactionist (inspired particularly by Mead) paradigms.[18] At the most general level, this proposal suggests that ontogenesis be understood as a form of progressive learning in which individuals become who they are through the acquisition and development of various interactive competences. Habermas's competence-theoretical model is thoroughly intersubjectivist in that it represents the process of individuation as occurring in and through processes of socialization into symbolically-structured social relations. The model identifies four types of generalized competence required for any person's meaningful actions in the world – epistemic, linguistic, normative-interactive, and motivational – as well as three types of competence contributing to the development of each person's distinctive sense-of-self as a particular individual – self-ascription, social role integration, and biographical integration.

On this model, learning is characterized as the decentering of the ego through expanding capabilities for demarcating the self from linguistic, natural, social, and subjective domains, for differentiating the respective symbolic logics of these domains, for generalization, for abstraction, and for reflexivity. The stability of one's ego is manifested in one's interactive consistency of speech and action at the level of competence achieved. Finally, a rich notion of autonomy is presented which characterizes it as an agent's rational accountability. An individual is fully autonomous when willing and prepared to defend before interlocutors her claims concerning, and her relations to, the various domains of language, the objective external world, the intersubjectively-shared social world, and the subjective world of inner nature.

---

[18] Besides the material in *Theory of Communicative Action,* see Jürgen Habermas "Stichworte zu einer Theorie der Sozialisation" and "Notizen zum Begriff der Rollenkompetenz" in *Kultur und Kritik. Verstreute Aufsätze* (Frankfurt: Suhrkamp, 1973), pp. 118–231, "Notizen zur Entwicklung der Interaktionskompetenz" and "Überlegungen zur Kommunikationspathologie," in *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns,* pp. 187–270, "Moralentwicklung und Ich–Identität" in *Zur Rekonstruktion des Historischen Materialismus* (Frankfurt: Suhrkamp, 1976), pp. 63–91, and "Individuierung durch Vergesellschaftung. Zu G.H. Meads Theorie der Subjektivität," in *Nachmetaphysisches Denken: Philosophische Aufsätze* (Frankfurt: Suhrkamp, 1988), pp.187–241.

Accordingly, an autonomous person – one with what Habermas calls a "post-conventional ego identity" – is prepared to defend, at ever higher levels of abstraction and reflexivity if need be, the meaningfulness and validity of her speech and actions as based upon defensible reasons rather than contingently given or taken for granted meanings, truths, conventions, values, etc.

Turning second to the phylogenetic analysis, Habermas contends that there are important homologies between stages of an individual's development and stages in progressive social development with respect to the cognitive structures of historical worldviews. In particular, he claims that we can retrospectively reconstruct sociocultural rationalization in a stage-sequential series of irreversible improvements in modes of consciousness that enable heightened problem solving through openness to discursive testing and rational belief fixation. In the domain of technical progress, this is largely a story of the increasing capacity for human's material control of the natural world made possible by revolutions in scientific and technical modes of thinking. But a parallel process of progressive rationalization can also be reconstructed in moral-practical domains, and precisely to the extent to which communicative action, rather than coercion or obedience to falsely naturalized authority, organizes ever more aspects of collective social life. Specifically, progressive modernization is understood as the rationalization of lifeworld structures: the differentiation of the various forms of validity claims, increasing abstraction from concrete relations, and ever increasing sociocultural dependence on risky processes of communication oriented to mutual understanding. Thus in the sphere of cultural reproduction, quasi-naturalistic traditions lose their security as increasingly reflective and abstract procedures are adopted for argumentation and knowledge acquisition. In the sphere of social integration, normative legitimacy is no longer secured through a particularistic, pre-given sacred ethos, but rather through formal discursive procedures for establishing and justifying generalized, reciprocal norms of action. In the sphere of individual socialization, identity formation is no longer fixed in concrete forms of life with a small set of required competences, but rather individuation must be self-directed in order to achieve autonomous self-realization and ego identities must become increasingly abstract in order to expand one's requisite competences.

This ambitious attempt to show that the standards of rationality celebrated in contemporary Western societies are not merely contingent conceits of a particularistic worldview, but can lay claim to universal, context-transcending validity is intended in part to answer to the various forms of skepticism raised against the ambitions of Enlightenment thinking. While Habermas develops his account of rationality in concert with the foil of Peter Winch's skeptical social scientific arguments inspired by Wittgenstein's account of particularistic language games, he also intends the account to answer to sociologists such as Weber, who contend that the various standards of rationality invoked by social actors are, in the end, no more than "warring gods and demons" with no justifiable principles of selection to appeal to. More importantly he intends this account to fill in a particularly glaring lacuna in the specific tradition of critical social theory: the inability of, for instance, Horkheimer, Adorno, and their colleagues to give a coherent justification for the ideals of individual autonomy, substantive social equality, and an emancipated society that they continually employed in critiquing the pathologies of sociocultural development. Yet one should not confuse Habermas's developmentalism with a Whiggish philosophy of history smugly justifying

the present or with a (right) Hegelian philosophical demonstration that the real is fully rational. The key here is to see that Habermas claims only to be reconstructing the *logic* of successive stages of consciousness, while making no parallel claims about the *dynamics* of historical development. Rather, the idea is that structural changes in the forms of consciousness can at most be understood as the pacemakers of history, but not the causal drivers of historical change. Social change is dependent on the contingent boundary conditions of the material reproduction of sociocultural life, and changes in these conditions are neither predictable nor developmentally progressive. In contrast to the claims of the later Marx, there are no iron laws of history dictating a systematic progression through various modes of production, but it is possible to reconstruct, at a suitable level of abstraction, socio-cultural learning processes that are both irreversible and clear improvements over earlier stages. Thus Habermas claims to be able to comprehend the normative content of modernity by employing a social theory built out of the elements of communicative action: a fallibilistic culture that is committed to critical testing of claims, a social solidarity founded on collective will formation through universal discourse, and personal socialization achieved though expanded individuation and self-realization.

The Integration of Systems Theory and Diagnoses of the Present

If the contingent dynamics of historical change are distinct from its progressive structural logic, then what explains the former? Here Habermas employs contemporary sociological functionalism to explain the reproduction of the material conditions of life. In essence, historical dynamics are to be understood in terms of responses to systemic steering problems encountered in the functionally integrated domains of the economy and state administration. To understand this, we need to look at a critical encounter in 1971 that was decisive for Habermas's mature social theory: his extended debate concerning the social systems theory that Niklas Luhmann developed by streamlining and radicalizing the functionalist theory of his teacher Parsons.[19] Although much of the debate turned on technical matters of sociological theory construction, at least two of Habermas's critical concerns are worth noting here. First, Habermas was convinced that no matter how powerful a functionalist perspective proved to be for illuminating economic and bureaucratic control processes, it grossly distorted phenomena when it turned to social domains that are fundamentally structured by language and the rich symbolic resources of communicative intersubjectivity. Reacting to Luhmann's radical strategy of analyzing meaning and truth themselves in functionalist terms, Habermas first pointed to internal problems in such a thorough-going systems theory. For instance, in order to say what the boundaries of a social system are, or what the target goals of a system are, systems theory would have to invoke meaningful definitions that are accessible only through the ordinary language of social participants. But since systems theory proposes avoiding ordinary language through a reductivist replacement by cybernetic codes, it cannot make the non-arbitrary boundary and goal-state definitions it requires. Systems theory seems to undercut even its own claims to be a truthful, accurate

---

[19] Jürgen Habermas, Niklas Luhmann, *Theorie der Gesellschaft oder Sozialtechnologie—Was leistet die Systemforschung* (Frankfurt: Suhrkamp Verlag, 1971). Several further follow-up volumes in the Suhrkamp series have been published containing papers from others on this topic.

representation of social reality to the extent that it understands "truth" as little more than a functional variable stabilizing a particular "cultural system" called science. The second main reservation Habermas had concerning systems theory was straightforwardly normative and political. While Parsons only implicitly endorsed limiting the scope of democratic action in favor of expert control over social subsystems in the name of efficiency and stability, Luhmann explicitly argued for de-democratizing various social spheres in order to take advantage of the supposed complexity-controlling achievements of technocrats schooled in systems theoretical analysis. He even went so far as to point out that this would require that many decisions be withdrawn from the explicit oversight of democratic politics and the public sphere, since stability often requires the "latency" – that is, the invisibility to social actors – of social structures and functions. In short, to the extent that a fully radicalized systems theory promotes a "counter-enlightenment" social technology, Habermas rejected the practical realization of functionalist insights in the name of the dialogical, public exercise of critical reason.

Looking back from the present, one can confidently say that no matter how forcefully and astutely Habermas critiqued Luhmann's systems theory at that time for its one-sided absolutization of the functionalist paradigm, it is clear that the encounter was very influential for his important 1973 sociotheoretic study of modern "steering crises" in economics and administration entitled *Legitimation Crisis*. Here Habermas put forward, in a programmatic and provisional way, a bold set of diagnostic theses concerning the interrelations, in contemporary capitalist democracies, between economic performance, administrative rationality, the extent of perceived legitimacy of the government, and the degree to which individuals are motivated to partake in business and politics. The basic thesis of the book is that crises in individual social subsystems are 'solved' by another subsystem, but then only at the cost of opening up that other subsystem to its own crisis potential. So, for instance the 20th century welfare state was able to solve economic crises of boom-bust cycles only at the cost of showing that government could not produce fully rational economic solutions. But to the extent to which administrative planning is economically effective, it also shows that the government works in the interests of capital rather than of all citizens, and thereby runs into legitimation crises. Such administrative and legitimation crises may finally ensue in a withdrawal of individual motivation in the face of alienation and malaise. Obviously the entire framework of this book is deeply indebted to systems theory, even as it holds out the hope that, in the end, norm and motive formation in contemporary societies are still constitutively tied to public, discursive justification, such that actual cultural meanings, social norms, and individual identities are not mere precipitates of functionalist imperatives, but respond to the independent logic of rational justification. Systems theory is also crucial to the final synthesis of functionalist and action-theoretic sociological perspectives achieved in the 1981 *Theory of Communicative Action*. In essence, Habermas did not just see the limits of systems theory, but was also persuaded by its power in illuminating the material reproduction processes of society and the causal mechanisms behind the tremendous growth, complexity, and success of contemporary capitalist economies and bureaucratic administrations.

The basic methodological idea behind the sociological synthesis in *Theory of Communicative Action* is to adopt a dual-perspective approach – lifeworld and systems – to investigating the coordination and integration of individual actions. The two

sociological perspectives are distinguished by the types of orientation taken up by participants in social action, the forms of coordination achieved, and the socially necessary reproductive functions performed in each action context. On the one hand, the lifeworld perspective attends to social actions oriented toward the task of reaching mutual understanding. Such communicative interactions are coordinated either by means of an already achieved store of background expectations and knowledge or by means of an explicit thematization of facts, norms, and values that is directed towards achieving consensus about them. Cultural reproduction is achieved through the communicative function of mutual understanding; social integration is achieved through the communicative function of coordinating action; and socialization is achieved through the communicative function of forming personal identities capable of autonomy and responsibility. On the other hand, the systems perspective attends to actions oriented towards the purposive achievement of actors' individual ends. In this case, interactions are coordinated by the functional imperatives of interconnected systems whose consequences are beyond the intentions of the actors. The systems perspective yields a functional account of economic and administrative subsystems, which are specialized in the material reproduction of society. The functional demands of increasingly complex forms of systemic integration lead to the delinguistification of action coordination in these subsystems, which are steered by the media of money and power. These media-steered domains are not directed by normative considerations of communication oriented towards mutual understanding and consensus formation, but are rather dominated by the logic of purposive-rational, strategic action.

With these two perspectives on action integration, critical social theory can isolate two interrelated phenomena characteristic of modernization processes. First, in modern societies, the rationality potential inherent in the structures of communicative action is realized through the progressive differentiation and rationalization of the three structural components of the lifeworld, already discussed above. Second, organizations through which systemic integration is achieved become uncoupled from the communicative relations in which they were previously anchored and, under the imperatives of ever increasing efficiency and self-maintenance, have grown into complexly interrelated subsystems capable of increased internal steering capacities. Thus, along with the progressive rationalization of the lifeworld in modernity, the media-steered subsystems also become increasingly differentiated internally, and their coordination mechanisms gain more independence from the lifeworld that makes their action coordinating functions possible in the first place. For, while the historical dynamics of social development are guided by imperatives of systems maintenance, these dynamics are constrained by the development and structural differentiation of lifeworld rationality structures which are achieved through communicative learning processes.

The critical moment of Habermas's project is already implied in this sociological reconstruction of western history. Processes of modernization can be seen, on this analysis, to be fundamentally *ambiguous:* both progressive and regressive. Insofar as the rationalization of the lifeworld releases the potential inherent in communicative action for reflective and relatively uncoerced discussion of, and agreement upon, facts, norms, values, and forms of life, modernity can be seen as overcoming the uncriticizable, merely factical normativity of traditional worldviews. Yet insofar as systemic forms of action coordination take over reproductive functions which can *only* be achieved through the

symbolic resources of the lifeworld, modernity surrenders essential decisions to functionally organized institutions which are steered by the value-free media of money and power, and whose consequences reach beyond the intentional plans and desires of individual actors. "When this tendency towards the uncoupling of system and lifeworld is depicted ... the irresistible irony of the world-historical process of enlightenment becomes evident: the rationalization of the lifeworld makes possible a heightening of systemic complexity, which becomes so hypertrophied that it unleashes system imperatives that burst the capacity of the lifeworld they instrumentalize."[20]

Finally, this analysis is also intended to identify the emancipatory content explicit or latent within current social practices. Critical social theory attempts to locate instances in which the functions of communicative action are taken over by media steered subsystems – that is, in which public and private spheres of life are relentlessly subjected to bureaucratic and economic imperatives – and thereby lead to a distortion and/or foreclosing of intersubjective consensual decision making. These instances become manifest either in steering crises in which the unintended consequences of systemic organization catch those organizations short, or when these steering crises are solved by instrumentalizing and thereby reifying the lifeworld resources of cultural reproduction, social integration, and socialization. In the latter case lifeworld *pathologies* result and conflicts arise, not around issues of material distribution, but rather in the three domains of lifeworld reproduction. The three main socially-caused pathologies Habermas identifies as due to the *colonization* of the lifeworld by systems are cultural loss of meaning (the assimilation of rich and meaningful ordinary language to the hollowed-out "semantics" of money and power), social anomie (the breakdown of integrating social norms and values), and individual psychopathologies (including withdrawal of motivation, disorienting senses of the loss of freedom, and mental illness). Although individuals may feel that the rationalization of the lifeworld itself leads to a loss of meaning and that the growth of media-steered subsystems of action leads to a loss of freedom, these transformations only become pathological, on Habermas's account, when systemic imperatives overextend their proper domain and colonize reproductive functions that can be achieved only through mutual understanding. By understanding the causes of these disturbances in lifeworld reproduction and by drawing on the emancipatory impetus embedded in communicative forms of sociation, social participants can then take up a reflective, critical stance towards pathological social consequences that had previously appeared as merely disconnected maladies of modern societies. With this grand synthesis of hermeneutic and systems theoretic approaches to sociology, Habermas claims to be able to better account for the social deformations that interested the great original sociologists of modernity – Marx, Durkheim, Weber, Lukács, Horkheimer and Adorno – without falling into the trap of a one-sided absolutization of one or another specific employment of reason.

Postmodernism and Poststructuralism

With this understanding of Habermas's ambiguous assessment of the promise and perils of modernization processes, we are in a better position to appreciate his somewhat orthogonal interaction with the concerns of poststructuralist and postmodern thinkers of the 1970s and 1980s. In his *Philosophical Discourse of Modernity,* concerned with the

---

[20] Habermas, *Theory of Communicative Action, Volume II,* p. 155.

distinctive sense of time-consciousness expressed in philosophical theories of modernity, Habermas seemed to finally acknowledge the continental and worldwide importance of the radical critique of reason that had developed in contemporary French thought under the influence of a distinctive reception of Nietzsche's and Heidegger's thought. The overall history of ideas thesis he pursues is that one should understand French poststructuralism as the culmination of a long-running critique of the philosophy of consciousness (alternatively labeled the philosophy of the subject), stemming from a rejection of Hegel's grand attempt to reconcile modernity with itself through absolute knowledge. Habermas dedicates chapters to many variants of this radical critique of reason: Nietzsche's and Heidegger's complementary destructions of subjectivist metaphysics, Horkheimer and Adorno's negative dialectics of instrumental reason becoming the central instrument of modern domination, Derrida and his American acolytes' deconstructionist transformation of philosophy into literature, Bataille's surrealistic celebration of the obscene, the impossible, and the taboo that have been constitutively excluded by ethical proscription, and Foucault's genealogical tracing of the deep but nevertheless historically contingent interconnections between tactics of modern power, the development of the human sciences, and the rise of contemporary subjectivity. In each case, the relentless critique of abstract, utilitarian Enlightenment reason and its corresponding figure of the sovereign, ratiocinating, decentered subject has been carried forward in the name of all of the impurities such conceptions of reason and the subject have left out: history, tradition, cultural specificity, power, desire, embodiment, rhetoric, metaphor, myth, narrative, ordinary practice, the unconscious, the irrational, the liminal, the non-identical, heterogeneity, contingency, idiosyncrasy, and so on. According to this radical critique, however, such impurities are not mere externalities, but are centrally and ineradicably constitutive of reason and subjectivity themselves.

Habermas's response to French poststructuralism and its Nietzschean forebears is basically two-fold. On the one hand, he agrees with their critique of the philosophy of the subject as a thoroughly exhausted philosophical paradigm that is doomed by both its internal contradictions and its idealizing disregard of the ineliminably situated character of reason and subjectivity. Yet on the other hand, he argues that the overly totalizing skeptical conclusions drawn from this critique are unwarranted. To begin, he argues that the radical critique of reason ends in its own aporias and paradoxes, particularly to the extent to which it leads to relativistic conclusions. Whether making truth claims that it cannot redeem in the face of its relentless critique of the very idea of truth, or relying in a cryptic way on normative intuitions about autonomy and nondomination while arguing that normative standards themselves are nothing more than effects of contingent relations of power and domination, the radical critique of the philosophy of the subject runs up against self-referentiality paradoxes. According to Habermas, however, there is an alternative path out of the correct critique of subject-centered philosophy: namely, the thoroughly intersubjectivist theory of communicative reason that sees reason and subjectivity as fully situated and immanent in everyday practices, but also as intrinsically oriented toward context-transcendence by virtue of their connections to validity claims. This alternative path can reinterpret the foundational ideas of the Enlightenment outside of the monological concepts that originally doomed their interpretation to the endless back and forth between subjectivism and objectivism. Instead of the epistemic ideal of the self-conscious individual subject gaining increasingly objective knowledge of the

external world, the theory of communicative reason points to intersubjective epistemic practices that allow disparate investigators to engage in collective learning processes through a fallibilistic attitude toward extant cultural contents and a commitment to critical testing; instead of the socio-political ideal of a self-determining macro-subject exhaustively organizing itself, the idea of social will formation through universal discourse aimed at increasing solidarity in an undamaged form of open intersubjectivity; and instead of the self-realizing sovereign individual subject undetermined by any externalities, the ideal of expanded possibilities for individuation and self-realization for intersubjectively socialized subjects. Finally, many of the specific socio-critical themes found in the earlier generation of critical theorists and in French poststructuralism – especially in the work of Foucault, to which Habermas's work nevertheless has many basic affinities – concerning the social dominance and power carried by instrumentalist and functionalist employments of reason are, according to Habermas, better understood in the ambiguous theory of modernization developed in *The Theory of Communicative Action* than in the totalizing critique of postmodernity.

IV   The Discourse Theory of Morality

One of the most significant components of Habermas's theory of communicative action concerns discourse theory – an account of the meaning of and justification procedures for the unconditional validity claims to truth, rightness, and sincerity that he claims are made at least implicitly in any communicative use of language. We have already briefly encountered his analysis of truth claims; it remains to look at his account of practical reason. (There is comparatively little that can be said philosophically about sincerity claims that, although they claim to hold unconditionally, can only be justified in the light of particular information about the consistency or lack thereof of the speaker's specific past behavior with purported truthfulness of their current claim). The central organizing principle of Habermas's normative theorizing – developed over the years in close connection with Apel – is summed up in what he calls the "principle of discourse," which specifies a general criterion for practically establishing normative validity: "Just those action norms are valid to which all possibly affected persons could agree as participants in rational discourses."[21] The central idea here is adopted from Mead and Kohlberg, but reflects the general idea that those affected by an action norm ought to be able to approve of it for themselves in order for it to gain their rational adherence. And the most plausible way of insuring this is to expose the proposed norm to public critical testing in the light of all relevant information, perspectives, and argument so that, in the end, an agreement concerning it can be expected to reflect only the weight of reason.

At this point, Habermas introduces the notion that there are different ways of employing practical reason, with claims of different scopes and types, and with distinct practical logics. *Pragmatic* questions concern the best means to adopt to some

---

[21] Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy,* trans. William Rehg (Cambridge, MA: The MIT Press, 1996), p. 107. Although the principles of discourse and of universalization were introduced and developed first in Jürgen Habermas, *Moralbewusstsein und kommunikatives Handeln* (Frankfurt, Suhrkamp, 1983), he revised both their specific formulations and, more importantly, his account of their relationship in the 1992 *Faktizität und Geltung,* and other later work. I refer throughout to the later conceptions.

contingently given preferences or goals; *ethical* questions arise when these preferences become problematic and one asks what is good for one to do in the light of who one is; *moral* questions arise when one's actions in pursuit of the good may conflict with the interests of others such that one must ask what universally applicable norms of action might govern anyone's actions in the same situation. The next move in the argument is to take the discourse principle, which is an intersubjectivist interpretation of the general idea of impartial justification, and see how it gets differently operationalized under different rules of argumentation for the different employments of practical reason. According to Habermas, applying the discourse principle to argumentation aimed at justifying moral norms generates the principle of universalization: "(U) For a [moral] norm to be valid, the consequences and side effects that its *general* observance can be expected to have for the satisfaction of the particular interests of *each* person affected must be such that *all* affected can accept them freely."[22] While the discourse principle can be operationalized in an appropriateness principle for moral application discourses and in a democratic principle for discourses justifying legal norms (discussed below), it can also be seen as operative in technical/strategic calculations concerning pragmatic questions, and even in the non-principled employment of reflective judgment operative in the hermeneutic and appropriative discourses concerning ethical-existential and ethical-political questions about the good, character, and identity.

Like Kant's categorical imperative or Rawls' original position, Habermas's (U) offers a procedure for testing the moral rightness or validity of proposed norms which meets four demands. It explicates the binding character of moral "ought" claims, remains at the level of formal procedures, depends on the cognitivist practice of giving reasons, and provides a universalist moral theory that transcends concrete forms of life. Unlike Kant's and Rawls' principles, however, (U) insists that the interests of those actually affected are morally relevant and thus that moral validity depends upon the real consensus of participants in actual discourses. It is this latter element of *intersubjectivity* at the heart of Habermas's theory that sets it apart from other impartialist moral theories. Discourse theory shows how the impartiality of the moral point of view can only be secured through actual reasoned dialogue, concerning concrete interests, amongst all those affected. Two other notable differences from traditional deontology are intended to insulate the discourse theory of morality from problems typically identified with it. First, (U) demands empirical assessments of the likely direct and indirect consequences of the observance of the norm, rather than trusting to either formal self-contradictions or unattractive possible worlds to ferret out norms that could not be universalized. Second, the consequences (U) considers concern individuals' *interests,* specifically those interests that they can convince others actually represent generalizable, even if not universally shared, interests. Thus (U) abstracts neither from the real world of consequences nor from the self-interpretation of the needs and interests of concretely-situated persons.

Whence the warrant for these ambitious claims concerning morality? Habermas combines a phenomenological account of the sense of normative obligation, a semantic account of the meaning of action norms, and a pragmatic analysis of the unavoidable presuppositions of engaging in practical justification (universal access to discourse, equal

---

[22] Jürgen Habermas, *Moral Consciousness and Communicative Action,* trans. Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, MA: The MIT Press, 1990), p. 120.

participation, non-coercion, decision on the weight of argument, and so on) in order to buttress his idealizing account of moral argumentation procedures. Moral philosophy, then, is another type of reconstructive science that attempts to elucidate the always-already presupposed, quasi-transcendental conditions that structure the actual moral discussions we already engage in – it focuses on moral communication as the immanent locus of the transcendent. Philosophy itself can only elucidate the formal nature of the procedures of moral argumentation, however; it has no special access to or claim over which putative norms are actually morally justified. For in the end, this is a matter for actual agreement amongst the universe of persons, and a philosopher is just another participant in the universal conversation. Notably, morality in this "post-metaphysical" view is not grounded in a transcendent being or reality, nor a particular ontological character or feature of the world, but rather is a thoroughly human, constructivist affair. There are no "facts of the matter" that operate as moral truth-makers. Thus unlike truth claims, ideal warranted assertability simply constitutes moral validity. Yet the conception is neither non-cognitivist nor decisionistic: some rules of action simply cannot be morally justified given the kinds of vulnerable creatures we are and the kinds of collective action problems we must solve. Meta-ethically, then, Habermas's position occupies a relatively rare position as a form of anti-realist moral cognitivism. Anthropologically, according to this view, morality comprises those norms that guard the fragile achievements of ego integrity which are always vulnerable to intersubjective misrecognition and harm, given that we only become individuated beings through socialization into a linguistically structured lifeworld. Hence moral practices constitutively involve the need for mutual recognition, reciprocal perspective-taking, listening to others' claims, a willingness to learn from others, and a responsiveness and responsibility to other's ultimate authority to agree or disagree with intersubjectively raised validity claims.

Of course universalist deontological approaches to morality (and closely-allied liberal approaches to justice) have come in for serious criticism during the same decades as Habermas developed his discourse theory; only the briefest indications of Habermas's extensive work in addressing such criticisms can be given here. First in response to the cultural or historicist relativist who claims that moral codes are no more than the contingent beliefs of determinate societies, it should be evident that Habermas's entire program is oriented toward rebutting relativism, mainly through a combination of the putatively trans-contextual empirical results of reconstructive sciences, the strength of the transcendental arguments concerning the pragmatic presuppositions of certain uses of language, the claim that such uses are anthropologically inescapable, and importantly, the restriction of the theory's claims to the formal and procedural features of moral discourses rather than the substantive first-order norms that different societies might believe to be justified. Next, in response to the radical moral skeptic who doubts that there is any cognitive content to moral claims, Habermas appropriates an argument developed by Apel to the effect that the skeptic must either engage in argumentation concerning morality – and thereby performatively presuppose the very standards s/he denies theoretically (a "performative self-contradiction") – or, on pain of psychopathology, withdraw from the socio-cultural form of life itself. Closer to home, Habermas is quite concerned to respond to a trio of structurally similar sets of objections (neo-Hegelian, neo-Aristotelian, and communitarian) that might be leveled against

discourse ethics for fundamentally distorting practical life in the same way Kantianism is often thought to: for being overly formalist, abstract, and disconnected from thick conceptions of who we are and who we want to be, for ignoring the qualitative character of the virtues, and for conceiving of moral agents as abstract, sovereign reasoners rather than real persons socialized into a shared ethical life and enmeshed in webs of personal and collective history and intimate relationships. Habermas's strategy here is not to facially rebut the criticisms, for he sees them as pointing to real truths about practical life: he repeatedly makes the Hegelian point that morality in the narrow sense requires an accommodating form of ethical life that anchors, fosters and sustains morality in cultural understandings, social interaction patterns and individual motivational structures. Rather he invokes the distinctions between moral, ethical-existential, ethical-political, and pragmatic employments of practical reason, agreeing that in real life ethical and pragmatic issues are often more pressing, salient and difficult to resolve than moral ones, but insists on the priority of the right over the good, that is, on the way in which a small set of universally binding moral norms puts constraints on our individual and collective pursuit of context specific ethical goals and values. Finally, in response to feminist care theorists and other moral particularists who critique abstract, rule-based moralities for their insensitivity to our thick commitments toward concrete others and the justified partial and non-symmetrical claims those commitments make on us, Habermas invokes the distinction between discourses of justification and discourses of moral application. Accordingly, relations of love, care, and concern that are responsive to concrete persons and their particular situations are to be grasped from the point of view of the principle of moral application which directs us to select that (presumptively justified) moral norm that is appropriate to the concrete situation in the light of a thick, exhaustive description of the particular situation.

V   The Discourse Theory of Law and Politics

1992 saw the publication of what might be considered Habermas's third magnum opus – *Faktizität und Geltung* – a book dedicated to a simple but extraordinarily ambitious thesis: "the rule of law cannot be had or maintained without radical democracy."[23] It brings all of the tools developed over the years to the domain that is arguably most central to Habermas's thought – politics – even though politics is investigated through its institutional infrastructure in modern nation-states: law. The reason for focusing on law is already announced in the title – literally *Facticity and Validity* – for modern, putatively legitimate law systematically presents a Janus face to the conflicting phenomena it simultaneously partakes in: factually structuring social relations in terms of ostensibly transcendent ideals, employing coercive force while drawing on convincing argument, constitutively shaping the subsystems of the economy and administrative power in its own esoteric code of legality while itself being shaped through the public use of practical reason in ordinary language, making binding claims on the actions of individual subjects that are ambiguously both factual and normative, and so on. Adopting the multidisciplinary, pluralist approach Habermas is famous for, the study combines four main analyses: a sociological and historical account of modern law and its social functions, a political philosophy reconstructing the normative validity claims made by constitutional democracy as the legitimate form of political rule, a dialogical

---

[23] Habermas, *Between Facts and Norms,* p. xlii.

jurisprudence aimed at comprehending the specific rationality of adjudication, and a normative-cum-empirical political theory explicating the ideals and realities of deliberative democratic politics.

From the historical-sociological point of view, the rise of modern positive law can be seen as a response both to the disintegration of medieval worldviews with their totalistic and encompassing religiously-cemented certainties and to the decoupling of economic structures from direct political (and clerical) control in the form of modern capitalism. For modern positive law makes direct, secular claims of authority over the actions of individuals through the threat of coercive sanction for nonperformance, while at the same time unburdening individuals from the constraints of communicative sociation according to norms in delimited spheres (notably the economic). Law thereby allows individuals to act rather as pure, purposive-rational actors calculating the costs and benefits of various courses of action. Yet modern law cannot be legitimated through its monopoly on the coercive use of force alone, for it also makes claims to being a rational, normatively correct structuring of social interaction, claims captured in liberal and republican social contract theories and manifestly informing the 18th and 19th century bourgeois revolutions. To put it in the terms of *Theorie des kommunikativen Handelns,* modern positive law partakes simultaneously in communicative and systemic forms of social integration – law "talks" in terms of both ordinary communicative language and the specialized codes of media-steered subsystems.[24] From the point of view of Habermas's critical social theory, this is a remarkable development, for now in place of rather vague references to potentials for social movement responses to manifold colonization problems in his previous work, he has now grasped an actually effective mechanism of social integration that bridges between the two other forms. In essence, law is now seen as being able to effectively employ communicative action to counter-steer functional subsystems run amok without however losing the apparently irreplaceable efficiencies of capitalism and rationalized bureaucracy for material reproduction. In the central metaphor of the book, law is the "transmission belt" which transforms social solidarity and mutual recognition into binding controls over anonymous, functionally integrated economic and administrative systems.

Showing how this sociologically effective medium of integration can be legitimate falls to Habermas's political-philosophical reconstruction of the social contract tradition, in particular of the normative core of constitutional democracy: the system of interlocked individual, political, and social rights, and the basic scheme for the separation of powers. These elements are interpreted in discourse-theoretic terms, such that the system of rights is grasped as exactly those rights individuals would need to legally grant one another if they wish to legitimately regulate their interactions through the medium of law, and the separation of powers is interpreted in terms of different ideal-typical employments of practical reason, where legislation has the function of justifying legal norms, adjudication has the function of applying them, and administration has the pragmatic function of administering them effectively. When the discourse principle is

---

[24] It should be noted that throughout the 1980s Habermas did not view law in these bridging terms, seeing modern positive law rather as itself a functionally integrated subsystem, with its own distinctive pathological form of the colonization of the lifeworld called "juridification."

operationalized in the domain of politics, it yields a principle of legitimacy for constitutional democracy: "only those statutes may claim legitimacy that can meet with the assent (*Zustimmung*) of all citizens in a discursive process of legislation that in turn has been legally constituted."[25] Once worked out, it becomes clear that this "principle of democracy" specifies a purely proceduralist understanding of legitimacy requiring real democratic participation of citizens in a multifaceted processing of reasons, opinions, and arguments, where that democratic deliberation is structured according to legitimate legal norms to ensure publicity, openness, and accessibility. Three distinctive features of Habermas's deliberative democratic account of constitutional democracy deserve mention. First, in distinction from classical and contemporary liberalism and their focus on the pre-political guarantees of private freedom, this theory decisively sides with the radical democratic side of classical republicanism in its endorsement of Rousseau's criterion for legitimacy: citizens can only understand themselves as free and equal if they are able to understand themselves simultaneously as the authors of the very laws they are subject to. But second, in accordance with the strict protections individuals would need to have if their political participation is to be procedurally secured from coercion and untoward influence, the discourse theory of democracy endorses strong legally entrenched individual rights, as well as political and social rights. According to Habermas, private and public autonomy are "equiprimordial": individuals can only enjoy equal individual liberties if they themselves deliberate and decide collectively about what is to be treated equally and what not. Said otherwise, the constitutional specification of rights must be attained through free and open democratic procedures, and constitution-making must be interpreted in fallibilistic terms as an on-going process of democratic deliberation by citizens about the basic terms of their legal consociation.

   This leads to the third distinctive feature of Habermas's account, which claims that democracy and constitutionalism are not antithetical ideals, but in fact mutually presuppose one another: democracy requires the rule of law to enforce procedurally required constraints, and the rule of law requires democracy to vindicate its inherent claim to legitimate legality. Habermas's jurisprudential analysis of law is focused around three debates important in German and United States legal academia. First he reinterprets Ronald Dworkin's theory of principled, common-law style adjudication in discourse-theoretic terms, attempting to show how the specific logic of legal discourse, in combination with the community of judges and legal scholars, yields the presumption of a heightened rationality from a court system charged with applying abstract norms to concrete situations. Second, he attempts to show how a strictly proceduralist understanding of constitutional adjudication can justify the otherwise surprising powers for constitutional legislation that are accorded to democratically unaccountable judges when they have the power (as they do in Germany and the U.S.) to rule on the constitutionality of legislation. He thereby attempts to vindicate the practice of judicial review against claims that it is fundamentally undemocratic. Third, agreeing with critical legal scholarship that points out how unacknowledged background paradigms of law's social roles systematically influence judicial decisions, Habermas nevertheless wishes to avoid drawing the skeptical conclusion that legal disagreements are thus no more than direct confrontations of competing paradigms and ideologies. He therefore develops a

---

[25] Ibid., p. 110.

social history of the 20th-century change in legal paradigms from a liberal to a welfare-state paradigm, and proposes the adoption of a "proceduralist" paradigm of law to adjust for the inadequacies of the welfare-state paradigm in dealing with the realities of the materialization of law.

The fourth major analysis of law in the book concerns the institutional means by which communicative power is politically transformed into administrative power. Notable here is the extension and modification of the model of the public sphere he developed in his *Habilitationsschrift* into a two-track model of politics. He now distinguishes between the informal public spheres of non-institutionalized, heterogeneous, and relatively anarchic arenas of debate and discussion found throughout civil society, and, the formal public sphere of state institutions oriented toward conclusive decisions justifying and applying legal rules, especially parliaments and legislatures but also including administrative and judicial bodies. Normatively speaking, there should be an ideal circuit of political power whereby, in response to problems identified and thematized in ordinary life, communicative power is first formed as public opinion in the informal public spheres, which then is worked up and formalized through the "sluice gates" of legislative processes, ensuing in legitimate legal rules that then direct the use of administrative power reacting back upon society. On this ideal picture, the circulation of power is only legitimated in terms of the robust deliberative character of opinion and will formation in both the informal and formal public spheres, since only that robust and open discussion warrants the expectation of rational outcomes from representative parliamentary procedures and thereby the legitimacy of the coercive use of state force. Of course, as Habermas recognizes, this is only an ideal, honored more in the breach. For the normal counter-circulation of power is from the center of the official government – itself often influenced by the social powers it itself is supposed to be steering through legitimate law – out to everyday society. Said another way, although modern law can ideally tame the excesses of functionalist subsystems through democratic means, the law making function is itself ordinarily instrumentalized to systems imperatives of money and power. Here one might say, the realist, social-scientific Habermas takes a step back from his revolutionary democratic convictions, arguing that, as long as ideal circuit of power can be put to use by a mobilized citizenry in times of heightened concern, the ordinary counter-circulation of power does not delegitimize the actual practices of contemporary constitutional democracies.

Through the 1990s and into the beginning of the 21st century, Habermas has applied the social-political theory of *Faktizität und Geltung* to any number of topics in both academic and broader public discussions: multiculturalism, collective identities, and social struggles for recognition; the future of nationalism and the possibility of a non-tribalist constitutional patriotism; tolerance in the liberal state between religious believers and nonbelievers; the status and character of political philosophy in a debate with John Rawls; citizenship rules and immigration policies; the justification of and prospects for the international extension of human rights; terrorism and the criminal law; increasing global inequalities and the ideologies and mechanisms of economic neo-liberalism; the changing face of international relations from the fall of the Berlin Wall to the unipolar moment of hegemony of the US, including a piece notably co-signed by Jacques Derrida arguing for a common European foreign and defense policy aligned with international

law as a counter-hegemon to the lawless, unilateralist US war machine.[26] This last topic of European unity has been central to Habermas's recent political writings, as he has argued for the adoption of a European constitution and the development of a European federalism that could realize the normative ideals of deliberative democratic constitutionalism on a transnational level. In a more sociological vein, he has also advocated the need for the Europe-wide development of the preconditions of successful democracy, including a collective European identity formed around Enlightenment ideals of freedom and solidarity rather than shared ethnic or religious origins and a set of public spheres of debate and discussion. He has also argued against a technocratic understanding of extant European political and regulatory institutions that would see their legitimacy largely as a function of specialized expertise leading to stability and efficiency.

Beyond this transnational level of Europe, Habermas has also devoted much attention to reconstructing and justifying the general outlines of Kant's cosmopolitan project for a supranational or global order. His argument here is that the ideals of constitutional democracy, *pace* Kant's substantive interpretation of cosmopolitan republicanism, are not best realized in a single global government but rather in the medium of law itself. Specifically, he argues for a constitutionalization of extant international law with an invigorated United Nations dedicated to securing human rights and promoting peace at the global level, while at the regional level transnational blocs would adopt various modes of federation, with democratic legitimation fed through the already exiting nation-state structures of constitutional democracy. In essence, this proposal seeks to radicalize an idea already in *Faktizität und Geltung*, namely, a desubstantialized, proceduralist understanding of the sovereignty of the people as no longer invested in a delimited set of citizens making up a bounded *demos*, but rather as resting in the very communicative structures and democratic procedures that allow for decisions to be made only in the light of sustained public criticism and testing and thereby warrant the expectation of heightened rationality as an outcome. Habermas seeks to operationalize this anonymous, subjectless conception of democratic sovereignty in national, transnational, and international governmental institutions and informal publics in order to realize the cosmopolitan project of peace and reconciliation through the rational taming of power by means of the constitutional rule of law.

VI   Systematic Philosophical Consolidation

In the last decade and a half, in addition to the ongoing outpouring of work on the broad themes treated throughout his critical social theory, Habermas has begun to publish various pieces that might be considered more in the domains of traditional philosophy,

---

[26] The piece cosigned by Derrida is Jürgen Habermas, "February 15, or: What Binds Europeans" trans. Ciaran Cronin, in *The Divided West* (Malden, MA: Polity Press, 2006), originally published in 2003. Most of Habermas's recent political writings can be found in the following collections in English: *The Inclusion of the Other: Studies in Political Theory,* ed. Ciaran Cronin and Pablo de Greiff (Cambridge, MA: The MIT Press, 1998), *The Postnational Constellation: Political Essays,* ed. and trans. Max Pensky (Cambridge, MA: The MIT Press, 2001), *Time of Transitions,* ed. Ciaran Cronin and Max Pensky (Cambridge, MA: The MIT Press, 2006), *The Divided West,* Ed. and Trans. Ciaran Cronin (Malden, MA: Polity Press, 2006).

specifically concerning ultimate questions of human meaning and concerning epistemology and metaphysics. Hence the last few years might be characterized as a kind of systematic philosophical consolidation, tying up various loose ends and addressing topics previously held slightly out of reach. In the domain of questions of ultimate human meaning, there are at least two quite surprising topics broached recently by Habermas. First, it is quite clear that for the last fifteen years, Habermas has written increasingly on religious and theological topics: appreciative essays on prominent theologians, interviews and articles treating the Christian and Jewish origins of ideas and thought complexes close to his work, and a more or less systematic reassessment of modernity as a project of Enlightenment. Notable here was an exchange with then-Cardinal Ratzinger (soon to become Pope Benedict XVI) in 2004 which contained an apparent shift in tone, if not wholly in substance, from his sociological theory of modernity developed two decades earlier. For while *The Theory of Communicative Action* couched modernization as a learning process involving the increasing abstraction, generalization, decentration, and differentiation of forms of thought – captured in the image of the structural differentiation and progressive rationalization of lifeworld structures of communicative action – it also couched these very same modernization processes in classical sociological terms: as "the disenchantment of religious-metaphysical worldviews," the "loss of the authority of the sacred canopy," and the "decentration of mythic and cosmological understandings of the world." In the Ratzinger exchange and elsewhere, however, Habermas is more sensitive to what has been lost with the changes in consciousness that he interprets as the linguistification of the sacred and as unambiguously leaving us in a post-metaphysical condition. Thus he now stresses that secular reason – which he still staunchly defends – must apply the canons of reflexivity to its own thinking, and open itself to a learning process whereby the irreplaceable symbolic and expressive potentials of religious experience are not wholly excluded. While consistently claiming that modernity has assimilated and reflexively transformed the normative content of religious traditions, these transformations have entailed diagnostic losses. Here Habermas refers "to adequately differentiated possibilities of expression and to sensitivities with regard to lives that have gone astray, with regard to societal pathologies, with regard to the failure of individuals' plans for their lives, and with regard to the deformation and disfigurement of the lives that people share with one another."[27]
In short, Habermas still sees secularized reason not only as unavoidable but also as justifiably superior to metaphysical and cosmological worldviews, but he is much more willing to consider and acknowledge the losses in semantic content and experiential sensitivity that religious disenchantment brings in its wake. Hence it is incumbent on believers and non-believers to open themselves to a two-way learning process in coming to understand the secularization of society. This idea is evident in his intervention into debates, spurred by Rawls, concerning the public, political use of religious reasons. In

---

[27] Jürgen Habermas, *The Dialectics of Secularization: On Reason and Religion,* with Joseph Ratzinger, trans. Brian McNeil (San Francisco: Ignatius Press, 2006), pp. 43–44. See also the following collections of Habermas's articles in English: *Religion and Rationality: Essays on Reason, God, and Modernity,* ed. Eduardo Mendieta (Cambridge, MA: The MIT Press, 2002) and *Between Naturalism and Religion,* trans. Ciaran Cronin (Malden, MA: Polity Press, forthcoming).

contrast to Rawls's endorsement of a restricted code of religiously cleansed "public reason," Habermas argues that religious reasons must be allowed in the informal political public spheres both for functional reasons and so that the special sensitivities of religious language for ethical deformations may be drawn on, as long as these reasons can be translated into secular reasons at the institutional threshold of the formal public spheres comprised of legislatures, administrations, and judiciaries.

The other important work on ultimate questions of human meaning concerns Habermas's intervention into bioethical debates, specifically concerning the ethics of liberal eugenics, that is, genetic interventions by potential parents aiming to improve or optimize their offspring in some way or another. Perhaps Habermas's conclusion is unremarkable – we should not engage in liberal eugenics – but the form of argument he puts forward is not only novel, but quite unexpected. In essence, he argues that various forms of genetic technology that might soon become technically possible would, if employed, fundamentally alter our species-wide self-understanding of ourselves as individual beings who are authors of our own lives and responsible for that authorship. With the knowledge that we have been *made*, to some extent, according to the optimization recipe decided upon by our parents rather than having *grown* according to nature and its fate, our sense of free agency is imperiled. But this massive change in our species-wide ethical-existential understanding of ourselves would then undermine our moral self-understandings as responsible authors of our own lives. Whether one shares the intuitions informing this argument, what is so remarkable is that, after decades of emphasizing the distinctions between the right and the good, between moral questions and ethical questions, between the context-transcendence of moral argumentation and the context-boundedness of ethical self-clarification, Habermas apparently here inverts their usual relationships. Universally valid moral norms are taken to be deeply embedded in a minimal but anthropologically universal species-wide ethical-existential self-understanding in such a way that the very bases of moral understanding would be shaken by the optimizing and self-instrumentalizing of the human race through genetic engineering. At the least, this work will force a reconsideration of central meta-ethical issues in Habermas's discourse theories of morality and ethics.

Habermas has also clarified and restated his views on questions of truth, language, objectivity, and reality as part of his systematic philosophical consolidation. As indicated above, in his most recent work Habermas is careful to differentiate truth and ideal warranted assertability in order to emphasize that, unlike in the case of morality and legitimate legal norms, truth is not constituted or exhausted by agreement under ideal conditions. There is always the possibility that the propositions we agree to even under ideal conditions could be false. This has lead to the recognition of a need for a theory of reference to complement this understanding of truth, and here Habermas has largely endorsed Hilary Putnam's theory of direct reference, thereby confirming what was always implicit: his epistemological realism. From the pragmatic point of view, our knowledge-gathering practices stem from problem-solving interventions in the world, interventions which make intersubjective learning processes possible through error correction and responsiveness to objections. It would be remarkable then, given the extent to which our experiences in such interventions can outstrip not only our expectations but also our linguistic frames of reference which provide our only access to the world, if we were to hold on to either anti-realist linguistic idealism or a hyper-

objectivist faith in direct access to brute empirical phenomena. In addition, he has sought to incorporate Robert Brandom's inferentialist semantics as a natural complement to universal pragmatics but, as in his treatment of other philosophers who focus exclusively on the representational use of language, Habermas has insisted that theoretical claims are not the only form of language use that needs to be conceptualized in a cognitivist semantic theory: language is involved first and foremost in communication, not just representation. Taking the linguistic turn in a fully pragmatic manner, requires one to articulate the connections between theoretical, normative, and aesthetic uses of language even as one carefully distinguishes between the different forms of justification through the intersubjective use of reasons appropriate to each. Finally, Habermas endorses what he calls a "weak naturalism" that treats both the natural and the sociocultural world as objective domains open to empirical investigation, yet insists against "strong" naturalism that one reduce the participant's perspective to pure causal analyses of socioculturally constituted facts. The normative self-understanding of competently speaking and acting subjects simply cannot be done justice to in a reformulated causal language of objectively observable events and states of affairs; the hermeneutic perspective is irreplaceable for comprehending the lifeworld, even as we need not thereby endorse hermeneutic idealism. In the end, Habermas continues to defend communicative practices as the immanent locus of the transcendent, now clarified for theoretical philosophy as a form of Kantian pragmatism acknowledging the constitutive contribution of historically-situated language to our grasp of the obdurately objective world revealed in experiences that go beyond our current linguistic resources and thereby spur us to reflexive transformations of those resources.