Pluralizing Universal "Man": The Legacy of Transcendentalism and Teleology in Habermas's Discourse Ethics

Stella Gaon

The central claim of this article is that Habermas's program of discourse ethics fails to "detranscendentalize" the Enlightenment subject. On the contrary, tacit assumptions concerning a transcendental conception of reason and a subject that is teleologically predisposed toward its rightful end are the logical pillars of Habermas's two most crucial claims. First, unless Habermas presupposes an abstract and decidedly unencumbered moral discussant, he cannot maintain his claim concerning the rationality—and hence the unconditionality—of the moral principle he describes. Secondly, unless Habermas begs the question of the proper end of individual and collective development, he fails to support the claim that discourse ethics speaks to the emphatic dimension of moral reason. Thus Habermas's formulation of discourse ethics does not overcome Enlightenment metaphysics; rather, its force depends upon the pluralization of universal "Man."

Introduction

Notwithstanding the fact that Jürgen Habermas's theory of discourse (or communicative) ethics has, over the years, provoked extensive and wide–ranging critique, the Habermasian approach to moral theory has largely remained intact. On one hand, a considerable body of research has been produced by those who stand in general support of the program—critics, that is to say, whose goal has been to raise and address narrowly specified problems within the theory and, thereby, to correct some of its more troublesome aspects with a view to strengthening the project as a whole.[1] On the other hand, of course, there are those who

1. See, among others, Albrecht Wellmer, "Ethics and Dialogue: Elements of Moral Judgement in Kant and Discourse Ethics," in *The Persistence of Modernity: Essays on Aesthetics, Ethics, and Postmodernism*, trans. by David Midgley

have challenged discourse ethics on a deeper level, particularly communitarians and neo–Aristotelians.[2] Yet so far these critics have been met with generally persuasive rebuttals. For example, in response to the common complaint that discourse ethics maintains an insupportable preference for the traditional subject of Western modernity, Habermas and his defenders have redirected attention away from discourse ethics' problematic dimensions to other moments of the theory, or have pointed to damaging weaknesses within the programs of communitarianism and neo–Aristotelianism themselves.[3] Such responses have been possible, I suggest, because strong opponents of discourse ethics have not produced a point–by–point, analytic demonstration of the way in which the program's unthematized suppositions undermine it at the most fundamental of levels.

(Cambridge, MA: The MIT Press, 1991); William Rehg, *Insight and Solidarity: The Discourse Ethics of Jürgen Habermas* (Berkeley: University of California Press, 1994); and Seyla Benhabib, "The Generalized and the Concrete Other: The Kohlberg–Gilligan Controversy and Feminist Theory," in *Feminism as Critique: On the Politics of Gender*, ed. Seyla Benhabib and Drucilla Cornell (Minneapolis: University of Minnesota Press, 1987), as well as Joseph Heath's more recent, "The Problem of Foundationalism in Habermas's discourse ethics," *Philosophy and Social Criticism* 21, no. 1 (1995): 77–100.

2. Peter Dews, for example, notes that the implication of Charles Taylor's critique is that "Habermas's whole conception of a discourse ethics ultimately rests on specific, albeit culturally deep–rooted, commitments to freedom and autonomy"—commitments which "cannot be derived from the normative structure of the speech–situation as such." See Dews's, "The Truth of the Subject: Language, Validity, and Transcendence in Lacan and Habermas," in *The Limits of Disenchantment: Essays on Contemporary European Philosophy* (London and New York: Verso, 1995), p. 274. Compare Charles Taylor, *Sources of the Self: The Making of the Modern Identity* (Cambridge, MA: Harvard University Press, 1989), p. 85. Compare also Christopher Zurn, who observes that William Rehg's contribution to communicative ethics implicitly privileges impartial reason over more contextualist forms of argumentation without adequately proving that there are no viable alternatives to the culturally specific, historical emergence of that post–Enlightenment ethos. See Zurn's "Review Essay: The intersubjective basis of morality," *Philosophy and Social Criticism* 22, no. 6 (1996): 115–19.

3. A case in point is Habermas's rebuttal to Charles Taylor's critique and his subsequent analysis of the shortcomings of Alasdair MacIntyre's approach. See Habermas's "Remarks on Discourse Ethics," in *Justification and Application: Remarks on Discourse Ethics*, trans. Ciaran P. Cronin (Cambridge, MA: The MIT Press, 1993), pp. 69–105.

For those who take issue with the universalist assumption that one version of moral theory can ultimately speak for us all, therefore, a crucial burden of proof remains. For this reason, my intention here is not merely to restate or to add to criticisms that have been produced before, nor is it to develop an alternative to Habermas's theory. Rather, my purpose is simply to show that discourse ethics is bolstered by a set of metaphysical assumptions which the theory itself cannot sustain. Indeed, I intend to demonstrate that when key critical points proffered by Habermas's most careful readers are brought forward together, with a view to exposing the metaphysical underpinnings of his overall approach, rather than independently, with a view only to rectifying or correcting specific issues on a case–by–case basis, these criticisms can be shown to constitute a decisive challenge to the discourse ethical program as a whole. What the following discussion shows, in particular, is that Habermas's reconstruction of "*the* moral point of view"—the perspective of impartiality and universality[4]— cannot avoid endorsing the Kantian subject; I argue that discourse ethics ultimately depends upon the transcendentalism and the *telos* of Kantian "Man." This metaphysical legacy, moreover, is structurally indispensable; it is embedded, ironically, in the very strategies through which Habermas distinguishes his position from that of Kant.

Among these strategies, the most important for my purposes is Habermas's proposed shift from solitary reflection to intersubjective agreement on moral norms. On this basis, in fact, Habermas explicates the moral "dignity" of our ability to universalize (*i.e.*, to reason impartially) *not* in terms of a noumenal Kingdom, but rather in the phenomenal terms of the intersubjective, hence *materially* or *socially* constituted nature of psychic integrity. In contrast to Kantian morality, then, discourse

4. Habermas defines "the" moral point of view as the *impartial* perspective in "Morality and Ethical Life: Does Hegel's Critique of Kant Apply to Discourse Ethics?" in *Moral Consciousness and Communicative Action*, trans. Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, MA: The MIT Press, 1990), p. 198. See also his "Lawrence Kohlberg and Neo–Aristotelianism," in *Justification and Application*, p. 118. For the sake of clarity, I retain Habermas's phrase ("the moral point of view") throughout this essay. The definite article is occasionally emphasized to problematize the coincidence between morality and impartiality assumed in this formulation.

ethics attempts to treat the reality of moral pluralism in a posttraditional world, and it does so precisely through a shift from monologic to dialogic modes of normative legitimation.

Now, once the actuality of moral pluralism is acknowledged, and once it is agreed that normative claims may be contingent upon one's social, historical, or cultural context, it becomes apparent that the possibility of universal moral validity depends directly on the difference between contingent and universal norms. In other words, if he is to de–transcendentalize Kantian moral theory by discarding the doctrine of the two realms without, at the same time, sacrificing the possibility of moral universalism, Habermas must draw a sharp distinction—here and now in everyday practice, so to speak; he must be able to distinguish between norms that are amenable to rational argumentation and therefore may be said to be universally justified—norms that constitute the so–called moral sphere—and norms that follow from particular notions of the good life and, as such, resist rational, consensual resolution (*i.e.*, so–called ethical norms).

My argument bears directly on this crucial demarcation. For, as Seyla Benhabib recently remarked, the division between justice and the good life is "truly an important point, and one to which sufficient attention has not been paid in the literature."[5] In contrast to Benhabib, however, the question that interests me most is not how discourse ethics might be reformulated *without* this sharp distinction in place, but rather why this divide is necessary and what, given its necessity, are its consequences. Specifically, what I want to underline is that Habermas's theory actually requires a morality–ethics split, and that this requirement gives rise, in turn, to two mutually contradictory necessities.

The problem can be spelled out as follows. First, Habermas must maintain the rationality of the moral sphere against incursions from the contingencies of ethical life. For it is only in this way that discourse ethics may claim to have justified a universalist perspective and, thereby, to have avoided falling prey to such biases as metaphysics or ethnocentrism. In other words,

---

5. Seyla Benhabib, "On Reconciliation and Respect, Justice and the Good Life," *Philosophy and Social Criticism* 23, no. 5 (1997): 104. Benhabib overviews the difference between the original Hegelian, and the contemporary Habermasian meaning of the terms morality and ethical life on pp. 103–104.

in light of the reality of moral pluralism in the postmetaphysical (posttraditional) world, the theory of discourse ethics must be conceived in purely procedural and formal terms, lest it be argued that, in being permeated by contingent, ethical goods, the moral sphere as explicated by Habermas reflects and imposes only the particular form of life which Habermas, among others, prefers.[6] In the second place, however, discourse ethics must still qualify as essentially *normative*, not merely analytic, in nature. For if Habermas insists that only a formalist ethic can meet the demand of rationality and thereby overcome the charge of cultural contingency, he is nonetheless aware that such an undertaking cannot lose its emphatic character altogether—that a purely procedural ethics risks being a trivial one, whereby it could be said that there is nothing substantively moral about it at all. To this extent, Habermas is also impelled to demonstrate the emphatic nature of moral universalism. Taken together, however, these necessities leave Habermas in the predicament of having to explicate *the* moral point of view in such a way as to render it *both* ethically empty *and yet* normatively full. Discourse ethics therefore will require compromises beyond what either of its aspects can bear.

6. Seyla Benhabib takes issue with this claim, arguing that there can be universal, as well as culturally–specific goods; she offers the example of the good of human rights. Yet Benhabib herself is very clear that, when conceived as a good rather than exclusively in terms of what is just, the ethos of human rights is in part culturally constituted and, as such, is not strictly rational; it is, therefore, *ethically* contestable relative to other goods. Indeed, based on the argument that a universalist morality of human rights and liberal tolerance follows from the perspective of the third person legislator, and that this perspective, in turn, entails culturally specific, albeit thin, notions of the good, Benhabib accuses Rainer Forst of coming closer in some respects "to an ethnocentric communitarian position" than she herself does, insofar as Forst would limit himself to that standpoint alone. Since it is this very charge of ethnocentricism that the autonomy of the moral sphere is intended to refute, however, such a softening of the line between morality and ethical life as Benhabib proposes will not serve Habermas's theoretical purposes. I return to this issue in section 2, below. See Benhabib's "On Reconciliation and Respect, Justice and the Good Life," esp. pp. 105–108, and Rainer Forst, "Situations of the Self: Reflections on Seyla Benhabib's Version of Critical Theory," *Philosophy and Social Criticism* 23, no. 5 (1997): 79–96. Peter Dews's solution to the problem of distinguishing sharply between justice and good coincides with Benhabib's. See his, "Morality, Ethics and 'Postmetaphysical Thinking'," in *Limits of Disenchantment*, p. 207.

Given the constitutive difficulty of this project, it should come as no surprise that Habermas's version of discourse ethics fails on both fronts. In what follows, therefore, I treat each aspect—the formal elucidation of communicative ethics and the argument for its normative bearing—in turn (sections 2 and 3). I show that in order to reconcile its contradictory ends, Habermas's theory must tacitly implicate both a transcendental notion of reason and a teleological conception of psychological maturation. As I shall argue, Habermas cannot discard this set of assumptions except at the cost of falling prey (in the first instance) to the charge of relativism and (in the second instance) to the charge of triviality. These analyses thus illustrate how, when and where the discourse ethics program shores up the metaphysics it putatively transcends.

On the Autonomy of the Moral Sphere

With regard to the first dimension of the Habermasian project, the attempt to address the pluralism of the modern world in postmetaphysical terms, we must begin by recalling that for Habermas, as for Weber, modernity is characterized by the fragmentation of reason into its different applications in the spheres of science, morality and art.[7] As a consequence of this historical result, Habermas claims, pre–Kantian concepts of substantive reason are simply not plausible; we no longer have recourse to "collectively binding religious or metaphysical worldviews."[8] This, quite simply, is what it means to Habermas to "take modern pluralism seriously."[9] Indeed, Habermas asserts, "only at the cost of Occidental rationalism itself could we rescind the differentiation of reason into those rationality complexes to which Kant's three critiques of reason refer. Nothing is further

---

7. See "A Reply to my Critics," in *Habermas: Critical Debates*, ed. by John Thompson and David Held (Cambridge, MA: The MIT Press, 1982), p. 240, and *The Philosophical Discourse of Modernity: Twelve Lectures*, trans. Frederick G. Lawrence (Cambridge, MA: The MIT Press, 1987), p. 339.

8. "Remarks on Discourse Ethics," p. 70. Habermas uses the phrase "historical result" specifically with reference to the development of moral universalism. See his, "Morality and Ethical Life," p. 208.

9. "Lawrence Kohlberg and Neo–Aristotelianism," p. 123.

from my intention than to make myself an advocate of such a regression, to conjure up the substantial unity of reason."[10]

It is on this basis that Habermas insists upon a narrow, purely procedural role for moral reason; given a plurality of worldviews, the theory of communicative ethics is intended to explicate only the *process of normative legitimation*, and is to be held entirely separate from evaluative questions of the good. "Moral theory is competent to clarify the moral point of view and justify its universality," he says, "but it can contribute nothing to answering the question 'Why be Moral?'"[11] Indeed, even more pointedly, Habermas says:

> What moral theory can do and should be trusted to do is to clarify the universal core of our moral intuitions and thereby to refute value scepticism. What it cannot do is make any kind of substantive contribution. By singling out a procedure of decision making, it seeks to make room for those involved, who must find answers on their own to the moral–practical issues that come at them, or are imposed on them, with objective historical force.[12]

Yet notwithstanding its evaluative neutrality, the formal delineation of communicative reason is intended to serve a crucial moral role. Specifically, it would not be unfair to characterize the driving motivation of the discourse–ethical project as a whole in terms of Habermas's desire to establish an incontrovertible moment of normative unconditionality for the theorization of social life. For Habermas, this moment is the *sine qua non* of a justified critique of society. Without the unconditional, he says,

---

10. "A Reply to my Critics," p. 235. And, even more strongly: "In contrast to the neo–Aristotelian position, discourse ethics is emphatically opposed to going back to a stage of philosophical thought prior to Kant" (see "Morality and Ethical Life," p. 206).

11. "Remarks on Discourse Ethics," p. 76. This particular essay is among Habermas's most recent and sustained attempts to clarify and correct his version of discourse ethics in light of critical objections from a variety of theoretical perspectives. It is therefore fair to say that this essay, along with, for example, "Discourse Ethics: Notes on a Program of Philosophical Justification" and "Morality and Ethical Life: Does Hegel's Critique of Kant Apply to Discourse Ethics?" (both in *Moral Consciousness*), represents Habermas's considered word on his position.

12. "Morality and Ethical Life," p. 211.

"we must be prepared to renounce the emancipatory potential of moral universalism and deny so much as the possibility of subjecting the structural violence inherent in social conditions characterized by latent exploitation and repression to an unstinting moral critique."[13]

Significantly, this insistence on the unconditional means that the distinction between "morality" and "ethical life" is absolutely crucial for Habermas's argument. In other words, if it is only by virtue of "the *moral* point of view" that one can arrive at agreements which are universal—that is to say, unconditional and therefore "rational"—in nature, then what Habermas calls "moral–practical discourses" do indeed "require a break with all of the unquestioned truths of an established, concrete ethical life"; they require, in fact, the "distancing [of] oneself from the contexts of life with which one's identity is inextricably woven."[14] Most importantly, to fail to determine a sphere comprised of universalizable norms would be to, "[succumb] to a relativism that robs moral commands of their meaning and moral obligations of their peculiar force."[15] Against Seyla Benhabib, then, we must concede that the morality–ethics divide is structurally indispensable; it concerns nothing less than the very basis, in Habermas's eyes, for a legitimate critical political response to social and political forms of injustice. The first issue at stake, therefore, is the strength of Habermas's case for the autonomy of the moral sphere.

With regard to this question, the thesis I wish to defend is that Habermas fails to distinguish clearly between the moral and the ethical spheres—that is to say, between "justice" and "the good"—except at the cost of burying a metaphysical premise in the theory of communicative ethics. In other words, I shall argue that dis-

13. See "Lawrence Kohlberg and Neo–Aristotelianism," p. 125. Similarly, Habermas elsewhere writes, "Even Marx set out his theory in such a way that he could perceive and take up the trial of reason in the deformations of class society. Had he not found in proletarian forms of life the distortion of a communicative form of life as such, *had he not seen in them an abuse of a universal interest reaching beyond the particular*, his analysis would have been robbed of the force of *justified* critique" ("A Reply to My Critics," p. 221, first emphasis mine).

14. "On the Pragmatic, the Ethical, and the Moral Employments of Practical Reason," in *Justification and Application*, p. 12.

15. "Remarks on Discourse Ethics," p. 76.

course theory is caught between the equally undesirable choices of either having to derive the legitimacy of *the* moral point of view from an unthematized reference to transcendence, or of falling prey to the charge of relativism Habermas has consistently sought to refute. Consequently, discourse ethics can be shown to privilege a subject that is removed, in essence and by definition, from the contingencies of everyday life.

This claim may seem strong, given that Habermas is fully aware that the insistence on universality carries a particular danger for contemporary deontologists. He knows, in particular, that Kant could appeal to the ideal of universalizability (in the form of the categorical imperative) as the very form of reason, only because the Enlightenment philosopher saw reason itself as the constitutive feature of the noumenal subject. Habermas, of course, no longer depends on such a strategy of argumentation; as he well knows, a postconventional, postmetaphysical perspective cannot invoke the subject's noumenal essence—what Kant called "the fact of reason"—to support the relation between reason and normative validity. At the same time, however, for Habermas the moral sphere only qualifies *as such* insofar as it is comprised of those norms which are amenable to rational justification through a procedure of universalization.

In order to solve this problem, Habermas proposes a reformulated principle of universalizability (U). For Habermas normative validity is conferred not merely by virtue of the fact that a moral actor employs the rational *ideal* of universalizability in a process of solitary reflection ("Act only according to that maxim whereby you can at the same time will that it should become a universal law"),[16] but rather by virtue of the *concrete enactment* of universalization in the form of an un–coerced consensus among a genuine plurality of communicative subjects.[17] Habermas formulates the principle (U), therefore, as follows: "a contested norm cannot meet with the consent of all of the participants in a practical discourse unless (U) holds, that is, "Unless all affected can *freely* accept the consequences and the

---

16. Immanuel Kant, *Grounding for the Metaphysics of Morals: On a Supposed Right to Lie because of Philanthropic Concerns*, third edition. Trans. James Ellington (Indianapolis: Hackett Publishing, 1993), Section II, par. 421.

17. See, "Remarks on Discourse Ethics," p. 51.

side effects that the *general* observance of a controversial norm can be expected to have for the satisfaction of the interests of *each individual*."[18]

Many theorists, of course, have drawn attention to the crucial shift Habermas proposes: from monologic to dialogic processes of justification. I shall return to this shortly. For the moment, however, I want to underline a particular aspect of this move. It is that, as Albrecht Wellmer has carefully argued, in order to qualify as valid, a norm must not only be considered through a process in which all interests are fairly and fully (reciprocally and symmetrically) represented—that is, a process in which all the reasons of all concerned are weighed equally and impartially so as to determine which reasons are the most forceful. Additionally, the norm must also be one to which all have *actually agreed*—that is, the interests of all concerned must also, in fact, have been served.[19]

This requirement becomes apparent once we notice, with Wellmer, that the justificatory process itself can only be redeemed (*i.e.*, determined as a valid one) *retroactively*. In other words, Wellmer argues, when taken as a principle of justice, Habermas's (U) is "quasi–circular": we can only say that all interests have been impartially represented in the first instance (*i.e.*, that all reasons and interests have been brought forward and equally weighted in the argument) in light of a consensus (*i.e.*, the determination of a norm which actually serves all interests) in the second instance.[20] Thus Habermas's explication of a common will escapes the charge that it is a metaphysical postulation precisely because it is a factual achievement; unless an actual agreement results from the discussion, there is no "common will" to speak of at all.[21]

18. "Discourse Ethics: Notes on a Program of Philosophical Justification," *Moral Consciousness*, p. 93.
19. Wellmer's argument is fully elaborated in "Ethics and Dialogue."
20. "Ethics and Dialogue," 149.
21. As Habermas writes, "As long as the isolated subject, in his role as custodian of the transcendental, arrogates to himself the authority to examine norms on behalf of all others, the difference between his supposition concerning a *general* will and an intersubjective agreement concerning a *common* will will never comes to light. . . . Once we abandon the metaphysical doctrine of two separate spheres of reality, subjects encounter each other as individuals who can no longer rely on [an] antecedent transcendental agreement" ("Remarks on Discourse Ethics," p. 51; compare "Morality and Ethical Life," p. 203).

It cannot be argued, therefore, that (U) provides only a formal presupposition *of*, or regulative idea *for*, argument—that, insofar as "communicative reason, unlike practical reason, is not itself a source of norms of right action,"[22] (U) is merely a rule of argument the *goal* of which is consensus. While this claim is fair as far as it goes, it obscures the fact that we cannot establish that a norm *is* legitimate unless we can establish that the principle of (U) has been applied, and—most crucially—*(U) cannot be said to have been applied unless an actual consensual agreement among impartial discussants has been reached.*[23] Significantly, therefore, while the conditions of the procedure itself are a necessary feature of its rationality (I return to this shortly), they are insufficient in themselves as the criterion of normative validity. Equally important is an actual consensus—the achievement of an agreement among us as to what our common (universal) interest really *is*—since only a consensus qualifies the strictly valid norm, and only the identification of such a norm, in turn, can retroactively establish the moral sphere. Indeed, it is precisely in this sense that Habermas speaks of the "discursive redemption" of validity claims.[24]

Once this is acknowledged, another significant difference between Kantian and Habermasian morality can be brought to light. In discourse ethics the meaning of universalizability shifts from what I (or we) can generalize *without contradiction* to what we have all (universally) *actually agreed* we *should* generalize.[25]

22. "Remarks on Discourse Ethics," p. 81, and "Discourse Ethics," p. 57f.

23. For Wellmer, the fact that discourse ethics requires an actual and not merely an ideal consensus means that Habermas is open to the charge that discourse ethics is nothing other than "the application of a general consensus theory of truth to the specific case of the concept of justice. To this extent (U_2) is not a specific principle of *justice* at all" ("Ethics and Dialogue," pp. 145–50, esp. 149). Wellmer's subscript here signifies that this is the second of four explications he offers for Habermas's principle (U). This explication reads, "(U_2) A norm is equally in the interests of all those affected precisely when it can be accepted without coercion by all those affected as being equally in the interests of all those affected."

24. "Remarks on Discourse Ethics," pp. 29, 51. For this reason above all it is clear that Wellmer's proposal that Habermas discard the requirement of consensus is simply implausible. See his "Ethics and Dialogue," p. 153f.

25. As Thomas McCarthy phrases it, "The emphasis shifts from what each can will without contradiction to be a general law, to what all can will *in agreement* to be a universal norm." Cited by Habermas in "Discourse Ethics," p. 67 (my emphasis).

Most importantly, this means that rationality can no longer be said to follow from the mere achievement of universality (*i.e.,* consensus)—as it did follow in the Kantian paradigm from the mere fact of being able to will without contradiction—since a universal agreement might be reached on any one of a variety of bases (agreement might be based purely on fear of reprisals, for example). The mere fact *that* we have agreed is not in and of itself evidence of rationality. Thus what Habermas calls the "rational potential inherent in everyday practice"[26] resides not in the criterion of universalizability alone but, additionally, in the *discursive process* whereby a universal agreement is achieved.[27] As Wellmer notes, the "structural characteristics of an ideal speech situation" become for Habermas the defining feature of rationality.[28] More specifically, an agreement is deemed fully rational only when it is based solely on the force of the better argument; namely, when the conditions *for* the agreement are (or at least adequately approximate) the ideal conditions of mutual recognition and reciprocity.[29]

Now this demand for an impartially–determined norm through actual consensus is precisely where Habermas runs into trouble. For it is precisely here that the definitive distinction between the categorically moral and the contingently evaluative must be established; here that an unconditional moral "ought" must be shown to be identifiable under real and not merely ideal discursive conditions, and here that the theory of discursive ethics must be identified as a moral theory in its own right rather than as a specific application of a consensus theory of truth applied to the realm of justice. Yet it is at this crucial moment that Habermas's

26. Habermas, *Philosophical Discourse of Modernity*, p. 341.
27. *Universal* here is to be taken to mean, "among all concerned with or affected by the norm in question"; that is, the rightful participants in the moral discourse.
28. See Wellmer, "Ethics and Dialogue," pp. 164–65 and 245, n.52.
29. Habermas qualifies the "redemption" of a validity claim in terms of "the framework of a discourse which is sufficiently close to the conditions of an ideal speech situation for the consensus aimed at by participants to be brought about solely through the force of the better argument, and in this sense to be 'rationally motivated'". Cited in Wellmer, "Ethics and Dialogue," p. 166. I leave aside the obvious objection that the "better" argument in moral disputes is rarely transparently evident.

theory of discourse ethics relies tacitly upon the possibility of transparent, ahistorical and culturally unencumbered rationality.

It is somewhat ironic that this conclusion can be demonstrated with reference to Seyla Benhabib's work, given her obvious sympathy with the Habermasian approach in general. But Benhabib has argued that the justice–ethical split that characterizes the onset of modernity translates into a split between the public and the domestic or private, with the consequence that the latter is simply left behind as part and parcel of an atemporal state of nature that is prior to history.[30] This means that the public, male figure of moral and political, justice–oriented theory is himself split "into the public person and the private individual"—a dualism Benhabib characterizes in terms of what she calls the "generalized" and the "concrete" other.[31] Most significantly, she suggests there is an "epistemic incoherence in universalistic moral theories" such as those of Lawrence Kohlberg and John Rawls, which entrench this dualism by focusing exclusively on the standpoint of the "generalized other."[32] She argues:

> We must ask whether the *identity* of any human self can be defined with reference to its capacity for agency alone. Identity does not refer to my potential for choice alone, but to the actuality of my choices, namely to how I as a finite, concrete, embodied individual, shape and fashion the circumstances of my birth and family, linguistic, cultural and gender identity into a coherent narrative that stands as my life story. . . . The self is not a thing, a substrate, but the protagonist of a life's tale. The

30. Benhabib, "The Generalized and the Concrete Other," p. 86. Significantly, while Benhabib is here treating Lawrence Kohlberg's theory in particular, there is no doubt that Habermas follows Kohlberg in this regard. He too characterizes the achievement of postconventional morality with its justice orientation—both socially and ontogenically—in terms of a catastrophic but natural break from a state of nature. See Habermas's "Moral Consciousness and Communicative Action," in *Moral Consciousness*, p. 126. I return to this in section 3 below.

31. The concept of a generalized other signifies the abstract, public persona of modern moral and political thought—it is based, notably, on the male head of the bourgeois household, and suggests the standpoint of *"formal equality* and *reciprocity."* The standpoint of the concrete other, in contrast, is based on the private world of personal and domestic life. This standpoint "requires us to view each and every rational being as an individual with a concrete history, identity, and affective–emotional constitution" ("The Generalized and the Concrete Other," p. 87).

32. *Ibid.*, pp. 88–89.

conceptions of selves who can be individuated prior to their moral ends is incoherent. We could not know if such a being was a human self, an angel, or the Holy Spirit.[33]

If the idea of a "self" prior to its concretization is incoherent as Benhabib suggests, and if, therefore, "there is no human *plurality* behind the veil of ignorance but only *definitional identity*, then this has consequences for criteria of reversibility and universalizability said to be constituents of the moral point of view. Definitional identity leads to *incomplete reversibility*."[34] Benhabib's argument thus suggests that the very distinction between justice and the good life which characterizes modern Western societies institutes a "generalized" concept of the self which renders universalistic moral theories constitutively incapable of accounting for concrete differences among actual moral discussants.

   This contention is directly relevant to the matter at hand, since Habermas, like Rawls, calls for "complete reversibility" in the specific sense that, "in communicative action speaker and hearer assume that their perspectives are interchangeable."[35] Indeed, we have seen that discourse ethical claim to rationality stands or falls with the impartiality of the process whereby consensual agreement is achieved. To this end, Habermas contends, at the level of argumentative praxis, "perspectives, relations of recognition, and normative expectations built into communicative action become completely reversible in all relevant respects, for participants in argumentation are credited with the ability to distance themselves temporarily from the normative spectrum of all existing forms of life."[36] There is a curious equivocation at work here, however, for unlike Rawls Habermas also insists that the role of such idealizations—for example, "crediting interlocutors with the ability to distance themselves from values of a evaluative, substantive kind"—does not entail the problem of empty formalism. On the contrary, he claims, "First, (U) regulates only argumentation among a plurality of participants; second, *it suggests the perspective of real–life argumentation*, in which all affected are admitted as participants. In this respect my universalization

33. *Ibid.*, p. 89.
34. *Ibid.*, p. 90.
35. "Lawrence Kohlberg and Neo–Aristotelianism," p. 131.
36. *Ibid.*

principle differs from the one John Rawls proposes."[37] Thus Habermas maintains that the idealization entailed by moral–practical discourse, the supposition that the conditions of the ideal speech situation sufficiently obtain,

> does not bear on the objects treated in argumentation; it leaves the identity of the participants and sources of conflict originating in the lifeworld untouched. The moral point of view calls for the extension and reversibility of interpretive perspectives so that alternative viewpoints and interest structures and differences in individual self–understandings and worldviews are not effaced but are given full play in discourse.[38]

This equivocation signals a crucial problem. If Habermas can be said to escape Benhabib's critique of Rawls by virtue of the fact that in his theory individual self–understandings and identities are not to be suspended but are, on the contrary, to be given "full play in discourse," then the procedure of moral discourse cannot be said to establish rationality by virtue of its impartiality. For in this case the strict criterion of impartiality will not have been met. On the other hand, if the individual participants in a moral discourse *are* fully to distance themselves from the contingent, "normative spectrum of all existing forms of life," the result for discourse ethics will be a transcendentalizing of the subject and the consequent "epistemic incoherence." Benhabib finds in moral theories that focus exclusively upon that standpoint. Either way, therefore, rationality cannot be said to be established solely by virtue of the discourse ethical procedure.

In response to Benhabib's analysis, Rainer Forst has recently argued that Rawls's description of an original position is intended only to establish the legitimacy of the principles of equality and reciprocity in discourse; "any further 'concrete' moral, political or legal questions," he says, "have to be dealt with in different ways," and "these ways obviously include others as concrete persons." To this extent, Forst argues, Kantian theories (such as

37. "Discourse Ethics: Notes on a Program of Philosophical Justification," p. 66; my emphasis.

38. "Remarks on Discourse Ethics," p. 58; emphasis mine. Elsewhere Habermas writes: "If actors do not bring with them, and into their discourse, *their* individual life–histories, *their* identities, *their* needs and wants, *their* traditions, memberships, and so forth, practical discourse would be robbed of all content" ("A Reply to My Critics," p. 255).

Rawls's) do not entail the sharp opposition between the "general" and the "concrete" that Benhabib describes.[39]

Such a critique, however—even if plausible—cannot be mobilized here. For there is a significant difference between Habermas's theory of justice and that of John Rawls, and to this extent Forst too quickly dismisses Benhabib's analysis as it bears on Habermas's work. Whereas as Rawls merely *posits* primary goods and an original position in order to arrive at principles of justice, Habermas seeks to *derive* those same principles from the structural features of communicative action as it might actually take place (*i.e.*, in "real life argumentation"). Importantly, this means that fully embodied, concretely–situated persons are necessary not only for the resolution of "further questions," but for the legitimation of the principles themselves. For this reason, it will not do for Habermasians to appeal to "an unavoidable 'ideal' moment in the sphere of morality," whereby the, "criteria of moral validity necessarily transcend[s] moral agreements . . . that have been reached."[40] For it is that very ideal, *qua* normative criterion, that Habermas is trying to establish in actual (pragmatic) rather than in transcendental terms.[41]

Habermas in fact foregoes Rawls's strong version of generalizing the other for precisely this reason; he insists that the legitimacy of norms *follows* from having given specific viewpoints and interests full play in a moral discourse.[42] In this case, therefore, it would appear that the discourse ethicist cannot do without the kind of concretization that Benhabib is calling for, since only concrete selves can engage in the full reversal of interpretive

39. Rainer Forst, "Situations of the Self: Reflections on Seyla Benhabib's Version of Critical Theory," pp. 93, 94.

40. *Ibid.*, p. 92.

41. See, for example, "Discourse Ethics," esp. pp. 75–76.

42. As Romand Coles notes, "In the idealizing supposition of a consensus open to criticism, the possibility of diverse voices on a given issue is not repressed, but rather the very condition of possibility for the legitimacy of the agreement" ("Identity and Difference in the Ethical Positions of Adorno and Habermas," in *The Cambridge Companion to Habermas*, ed. Stephen K. White [New York: Cambridge University Press, 1995], p. 25). Compare also Habermas's own comment: "Discourse ethics prefers to view shared understanding about the generalizability of interests as the *result* of an intersubjectively mounted *public discourse*" ("Morality and Ethical Life," p. 203).

structures that Habermas demands. Consequently, however, he must give up the discourse ethical claim to impartiality as well. For concretely–situated others, in contrast to general or abstract ones, are decidedly partial, precisely to the extent that they *cannot* be said to be removed from the normative spectrum of their everyday lives. The implication of Benhabib's argument for Habermasian discourse ethics is thus that the impartiality of the moral point of view is bought at the cost of transcendentalizing the subject. On the other hand, insofar as Habermas intends a genuine and not merely a nominal pluralization of moral interlocutors, his distinction between morality and ethical life— the very distinction upon which the rationality of discourse ethics depends—is untenable. Indeed, Benhabib notices that it was the very shift to a justice or moral orientation within modern political thought that gave rise to the general–concrete difference in the first place.

Notwithstanding this difficulty, Habermas is quite right to take issue with Benhabib's own solution. Her proposal is to replace a strictly moral theory with a "communicative ethic of need inter-pretation," whereby "the object domain of moral theory is so enlarged that not only rights but needs, not only justice but pos-sible modes of the good life, are moved into an anticipatory–utopian perspective."[43] Yet once it is agreed that evaluative or ethical goods are necessarily implicated in justifica-tory processes of moral legitimation, the discourse ethics project is undone altogether. As Charles Taylor explains:

> The boundary between questions of ethics [*i.e.*, Habermasian "morality"], which have to do with interpersonal justice, and those of the good life [*i.e.*, "ethics"] is supremely important, because it is the boundary between

---

43. Benhabib, "The Generalized and the Concrete Other," p. 93. See also her "Afterward: Communicative Ethics and Current Controversies in Practical Philosophy," in *The Communicative Ethics Controversy*, ed. Seyla Benhabib and Fred Dallmayr (Cambridge, MA: The MIT Press, 1990), pp. 349, 350. Forst notes that this move seems to reflect an Hegelian–influenced vision of reconciled life which is not itself supported by the theory. See his "Situations of the Self," esp. pp. 79–81, as well as Benhabib's response, "On Reconciliation and Respect, Justice and the Good Life." For further criticism of Benhabib's suggestion, see also Herta Nagl–Docekal's "Seyla Benhabib and the Radical Future of the Enlightenment," *Philosophy and Social Criticism* 23, no. 5 (1997): esp. pp. 70–71.

demands of truly universal validity and goods which will differ from
culture to culture. This distinction is the only bulwark, in Habermas's
eyes, against chauvinistic and ethnocentric aggression in the name of
one's way of life, or tradition, or culture. It is thus crucial to maintain it.[44]

Clearly, Habermas must reject Benhabib's version of discourse
ethics; his refutation of the charge of relativism rests directly on
the contention that the universal core of our normative intuitions
can be distinguished absolutely from the contingencies of ethical
life on the basis of the application of the principle of
universalizability (U).

Now in Habermas's reconstruction of (U), it will be recalled,
the rationality of moral action shifts from its basis in the Kantian
injunction to universalize without contradiction to that of a pro-
cedure aimed at a universal consensus guided by impartial
reasoning and ideal role–taking. Yet we have also seen that the
impartial reasoning and ideal role–taking Habermas has in mind
are the characteristics of a generalized other that is definitionally
incapable of undertaking the full reversibility upon which
Habermas himself insists. Indeed, the principle (U) may be said
to have been applied only when concretely situated individuals
(each of whom necessarily brings his or her substantive, ethical
concerns and perspectives into the discussion) have actually par-
ticipated in a debate. Consequently, no purely rational,
non–relative, unconditioned sphere of morality can be said to
emerge *from* the procedure itself.

Without recourse to an unthematized transcendentalism, then,
(U) cannot be relied upon to mark *retroactively* the distinction
between questions of justice and questions of the good life and,
thereby, to confer moral validity (in the form of unconditionality)
on the process of normative legitimation. Such unconditionality
would require a *prior* distinction between morality and ethical
life; without a distinction between justice and the good *already* in
place, there is no basis for the postulation of a generalized
(transcendental) other capable of the pure impartiality moral
legitimacy requires. Without the prior availability of this
distinction, in other words, we must face the reality of concrete
discussants who are not impartial, transcendental subjects, but

44. Taylor, *Sources of the Self*, p. 88.

are situated, partial individuals whose moral arguments are inextricably layered with contingent evaluations. Consequently, the theory of discourse ethics must either fall prey to the charge of relativism, or it must continue to rely on an abstract notion of the human subject—it must rely on a subject, that is, whose transcendence of the historical, gender, racial, class, and cultural (*inter alia*) contingencies of everyday life is not merely a temporary achievement but a constitutive attribute.

Habermas's attempts to address Benhabib and Gilligan's critiques of the Kohlbergian self do not resolve this difficulty with specific regard to moral justification. On the contrary, he shifts the terms of the debate, arguing:

> Practical reason is not fully realized in discourses of justification. Whereas in justifying norms practical reason finds expression in the principle of universalization, in the application of norms, it takes the form of a principle of appropriateness. Once we grasp the complementarity of justification and application, it becomes clear how discourse ethics can address the misgivings you [Torben Hviid Nielsen] share with Seyla Benhabib and Carol Gilligan.[45]

Not only does this response leave the problem of transcendentalizing the subject intact with regard to justification, but it brings to light a further issue. That is, upon examination it emerges that the determination of appropriateness (A) in the context of application no more reflects a contextual mode of thought than does the determination of universalizability in the context of justification. Rather, just as (U) entails abstracting from normative worldviews, so (A) requires that, "An *impartial judge* must assess which of the competing norms of action—whose validity has been established in advance—is most appropriate to a given concrete case once all of the relevant features of the given constellation of circumstances have been accorded due weight in the situational description."[46] Leaving aside the arguable assertion

45. Interview with Torben Hviid Nielsen in "Morality, Society, and Ethics," *Justification and Application*, p. 154.

46. "Lawrence Kohlberg and Neo–Aristotelianism," p. 128f., esp. pp. 129–130 (my emphasis). On the principle of appropriateness, this would seem to be Albrecht Wellmer's view as well. See his "Ethics and Dialogue," pp. 202, 203, wherein it emerges that, insofar as Wellmer presupposes there can be only one

that "the" relevant features of a moral question could be definitively determined, we need only note that the original charge may be redirected to the issue of application as well: if the judge of appropriateness is *also* to be impartial (by distancing him or herself from contingent notions of the good), it is not clear either how that achievement is to be determined, or in what way, precisely, the necessity of concretizing moral discussants in their socio–historical contexts has actually been addressed.

It would certainly seem, then, that the program of discourse ethics is caught at an inescapable impasse regarding the demand to distinguish between questions of justice and questions of the ethically good life. As the argument concerning the partiality of moral discussants bears out, Habermas must either concede the relativism of discourse ethics, or he must support the autonomy of the moral sphere on the basis of a metaphysically determined concept of pure or transcendental reason. In either event, the demonstration of a *post*metaphysical break between questions of the good and questions of justice has not been provided; rather, the neutrality of the discourse ethics program relies upon an unacknowledged metaphysical stake.

On the Emphatic Nature of Moral Unconditionality

So far I have been arguing that Habermas fails to demonstrate the autonomous status of the moral domain vis–à–vis what he calls conditional, evaluative questions of the good. Yet crucial though this problem is, it is still only one dimension of what I have characterized as an essentially dichotomous project. That is

correct interpretation of a given situation, questions of appropriateness raised by the application of norms in concrete situations lend themselves to the moral point of view. Compare Habermas's "Moral Consciousness and Communicative Action," in which he claims, "Interpreted from the perspective of discourse ethics, practical reason does indeed require practical prudence in the application of rules. But use of this capacity does not restrict practical reason to the parameters of a specific culture or historical period. *Learning processes governed by the universalistic substance of the norm being applied* are possible even in the dimension of application" (pp. 181–82; emphasis mine). If I understand him correctly, Habermas is here saying, once more, that the prudential question of appropriateness can be answered impartially—eventually—as well. This interpretation is supported in the interview with Hviid Nielsen ("Morality, Society, and Ethics," p. 172).

to say, one may, on one hand, accept the contradiction inherent in the first aspect of Habermas's project by acknowledging that communicative ethics achieves its status of unconditionality by virtue of implicating an unthematized transcendentalism. Lacking a transcendental subject, the proceduralism in question may be seen to be culturally and historically loaded; it is not, however, liable to the charge of triviality since, on this view, substantive, ethical beliefs are always already implicated in moral discourses.

On the other hand, if one can establish the autonomy of the moral sphere by resolving the problem characterized here as Habermas's pluralization of transcendental man, the possible triviality of Habermas's program becomes a real risk. For if (U) *could* truly be said to act "like a knife that makes razor sharp cuts between evaluative statements and strictly normative ones"[47]—if the explication of (U) *had* actually established a definitive separation between the moral sphere and ethical life—then Habermas would confront an additional and perhaps impossible task. In the face of a definitive divide, he would have to substantiate his own crucial contention "that the *procedural* explanation discourse ethics gives of the moral point of view—in other words, of the impartiality of moral judgments—constitutes an adequate account of moral intuitions which are after all *substantive* in kind."[48]

In order to counter the charge of triviality, then, Habermas must demonstrate a necessary connection between a procedur-ally–derived universalism, and the substantive nature of moral belief; he must somehow reunite—*without compromising its ethical neutrality*—the sphere of moral validity with that of evaluative goods. Yet the theory of discourse ethics, I shall argue, does not adequately support the possibility of an unconditional feature of ethical life. Rather, the unconditional normativity of the Habermasian "ought" is on my view conditioned by an unthematized teleology.

Given my contention, it is important to note that, on the face of it, Habermas is very clear that a truly postmetaphysical, deontological theory can provide no answer to the question "Why be Moral?": "Once the bond between the right and the good is

47. See "Discourse Ethics," p. 104.
48. "Morality and Ethical Life," p. 199.

broken," he writes, "the question of why one should act morally at all can no longer be answered satisfactorily."[49] Nor can moral philosophy awaken moral perception, for moral action is affectively impressed upon us "prior to all philosophizing."[50] Indeed, in marked contrast to what he sees as the teleological moment of Karl–Otto Apel's version of discourse ethics, Habermas insists that "communicative reason is not itself a source of norms of right action" and, to this extent, it provides "only weak rational motivations" rather than a "binding practical orientation."[51]

Yet this is not to say that Habermas views discourse ethics as no more than a normatively empty procedural mechanism. On the contrary, he speaks regularly of "the *substantive* normative presuppositions of argumentation"[52] which the theory is intended to elucidate. In fact, we have seen, what is at stake here is nothing less than, "the possibility of subjecting the structural violence inherent in social conditions . . . to an unstinting moral critique."[53] Now if the procedural rationality characteristic of modernity is intended not only, as Habermas says, "to give credence to our views in the area of moral–practical insight",[54] but indeed, to serve as the very basis of a justified critique of society, then Habermasian moral theory must do more than specify "to each discussant an equal and reciprocal share." Additionally, the theory of discourse ethics must explicate the precise sense in which such procedural necessities as (U) might command are *substantively* normative in nature.

This demand cannot be met merely with reference to the rational dimension of (U) as it has been elucidated so far. As Wellmer insightfully argues, even if (U) could be said to ensure the rationality of the moral domain by giving rise to genuinely universal claims, this feature of discourse ethics alone would not suffice to explicate a specifically normative "ought." For, most importantly, the corollary of replacing Kant's metaphysical doctrine of a Kingdom of Ends with the process and outcome of

49. "Lawrence Kohlberg and Neo–Aristotelianism," p. 119.
50. "Remarks on Discourse Ethics," pp. 75–76.
51. *Ibid.*, p. 81.
52. See, for example, "Remarks on Discourse Ethics," p. 83 (emphasis mine).
53. "Lawrence Kohlberg and Neo–Aristotelianism," p. 125.
54. "Philosophy as Stand–In and Interpreter," in *Moral Consciousness*, pp. 3–4.

real argumentation is not merely that we no longer merely *postulate* universality. It is also that we can no longer tie the force of reason's demand to our self–respect or dignity as supersensible beings.

Clearly, Habermasian "morality" no longer signifies the self–respect that accrues to noumenal subjects; for Habermas the term *moral* refers instead only to the much thinner notion of universalizability. In this sense, Wellmer contends, Habermas collapses two distinct features of the Kantian imperative. For Kant, it should be recalled, acting in contradiction of the principle of reason would be acting, "in contradiction of the conditions for the possibility of our self–respect *as* rational beings."[55] Yet here the *sense* and the *condition* of morality are not the same thing: since Kant already knows what morally–right action *is* (*i.e.*, action in accordance with our nature as supersensible beings), he addresses only the condition of its realization (*i.e.*, the postulation of a Kingdom of Ends.)[56] In Habermas's reconstruction, on the other hand, (U) is intended to serve as both the condition *and* the sense of morality; (U) is both a rule *for* the legitimation of moral norms, and the meaning or sense *of* those norms themselves in terms of what Wellmer calls an "elementary" concept of justice as equality.

Yet even in Habermas, Wellmer argues, the distinction should be maintained. The generalization (or equality) principle is, strictly speaking, merely a logical principle of legitimation; it tells us to be consistent in our behavior by treating like cases equally, and, on this basis, allows us to determine whether a norm meets the conditions for legitimacy. However, insofar as the elementary concept of justice (equality) that this rule entails is—at least on Wellmer and Habermas's view—the only normative concept which could *itself* be legitimized on these terms (*i.e.*, consensually agreed to by everyone concerned), the sense of morally right action and the conditions for its realization seem to coincide: *both* are determined as "universalizability" in Habermas's moral theory, and in this sense, a more substantive sense of "morally right action" is accordingly lost.

In fact, once morality has been reformulated as Habermas proposes, it is no longer obvious that (U) conveys substantive normative content. For if no argument is made specifically

55. "Ethics and Dialogue," pp. 151–152.
56. *Ibid.*, pp. 120–21, 122.

concerning its normative dimension, (U) could be fairly said to represent a general principle of distributive justice (*i.e.*, justice as participatory equality), but not a strictly moral principle. Given Habermas's divergence from Kant, then, he is faced with the added burden of providing a nonmetaphysical explanation of the relation between rationality and normativity. Unless he can do so, discourse ethics will amount to no more than an amoral explication of the purely logical imperative not to contradict ourselves. In this sense, the necessity at stake is that of being able to show not only that agreements reached among discussants are rational —as might be said of a consensual agreement on *any* question— but that they are also necessarily moral, which is to say normative, in nature.

It is to this end that Habermas insists on the emphatic dimension of communicative ethics. Indeed, he directly addresses charges of empty formalism and abstract universalism by arguing that participants in practical discourse are inevitably exhorted to partake in ideal role taking—that is, to accept the normative content inherent in the pragmatic presuppositions of argumentation—and that participants are "*constrained* to speak and act under idealized conditions."[57] Since the exhortation and constraint we experience as communicative agents is directly consequent upon our intersubjectively (hence collectively) constituted identities, Habermas argues, discourse ethics entails an ineradicable concern for the common weal. As he says, "though organized around a concept of procedure, [discourse ethics] can be expected to say something relevant about substance as well and, more important perhaps, about the hidden link between justice and the common good."[58] And it is in making *this* case, I contend, that a teleological assumption is unavoidably implicated in the communicative ethics project.

To see this we must note that, for Habermas, (U) is derived from procedural rules of argumentation—that is, from a reflective form of communicative action—and communicative action in turn entails both a strategic and an understanding orientation. Thus Habermas attempts to demonstrate that a bridging principle which makes consensus possible—the formal principle (U)—can

57. "Morality and Ethical Life," pp. 198, 203.
58. *Ibid.*, p. 202.

be derived from the very rules of discourse we already accept, "in conjunction with a 'weak' idea of normative justification."[59] He says:

> There is no form of sociocultural life that is not at least implicitly geared to maintaining communicative action by means of argument, be that actual form of argument ever so rudimentary and the institutionalization of discursive consensus building ever so inchoate. Once argumentation is conceived as a special form of rule–governed interaction, it reveals itself to be a reflective form of action oriented toward reaching an understanding. Argumentation derives the pragmatic presuppositions we found at the procedural level from the presuppositions of communicative action. The reciprocities undergirding the mutual recognition of competent subjects are already built into action oriented toward reaching an understanding, the action in which argumentation is *rooted*.[60]

Most importantly, "that is why the radical sceptic's refusal to argue is an empty gesture": on the basis of the quasi–transcendental demonstration that this refusal entails a performative contradiction, Habermas contends that the sceptic "remains bound" to the presuppositions already implicit in the "communicative practice of everyday life."[61]

As I have already suggested, our being bound to the communicatively derived presuppositions of argumentation is crucial to the task of explaining the normative dimension (as opposed to the rationality) of a specifically "moral" domain. By rooting a virtually inescapable rule of argumentation in action oriented toward reaching an understanding, therefore, Habermas seeks to ground the *normative* ideals of reciprocity and symmetry (*i.e.*, the impartial point of view) as the sense—the obligatory force—of the moral "ought."

We can now understand, I think, Wellmer's observation that, "what is expressed in the unconditional character of the moral 'ought' is the fact that our possible identity as creatures capable of speech is tied to . . . a structure of intersubjectivity."[62] For as

---

59. See "Discourse Ethics: Notes on a Program of Philosophical Justification," pp. 68, 92.

   60. *Ibid.*, p. 100.

   61. *Ibid.*, pp. 100–101.

   62. "Ethics and Dialogue," 152.

Habermas explains, morality can be understood in "anthropological terms" as a "safety device compensating for a vulnerability built into the sociocultural form of life." He maintains as a "basic fact" that we are constituted as subjects at all only insofar as we 'externalize' ourselves by engaging in intersubjective communication, and that this engagement results in "an almost constitutional insecurity and chronic fragility of personal identity."[63] Similarly, he says, "morality is aimed at the chronic susceptibility of personal integrity implicit in the structure of linguistically mediated interactions, which is more deep-seated than the tangible vulnerability of bodily integrity, though connected with it."[64] To the extent that our very possibility as subjects depends upon our communicative interaction, then—and indeed Habermas insists that the alternatives to such interaction are schizophrenia and suicide in the long term[65]—the impartiality that characterizes the moral point of view as elucidated within discourse ethics is a quasi-natural, inescapable, *affectively* compelling aim of both individual and social development.

This is why, I want to propose, Habermas relies—as indeed he must continue to rely—on the developmental psychology of Lawrence Kohlberg.[66] For against the charge of relativism,

> Kohlberg's theory of moral development offers the possibility of (a) reducing the empirical diversity of existing moral views to variation in the *contents*, in contrast to universal *forms*, of moral judgment and (b) *explaining the remaining structural differences between moralities as differences in the stage of development* of the capacity for moral judgment.[67]

63. "Morality and Ethical Life," p. 199. On this point, see also Habermas's "Moral Development and Ego Identity," in *Communication and the Evolution of Society*, trans. and intro. Thomas McCarthy (Boston: Beacon Press, 1979).

64. "Remarks on Discourse Ethics," p. 109.

65. "Discourse Ethics," pp. 102, 100.

66. Habermas says that "the normative reference point of the developmental path that Kohlberg empirically analyzes is a principled morality in which we can recognize the main features of discourse ethics" ("Moral Consciousness and Communicative Action," p. 117).

67. *Ibid*; last emphasis mine. Significantly, the claim that all moral views can be contained within the same formal structure holds only insofar as Kohlberg can be said to have achieved a plausible account of *human*—and not just western, middle class, white, male—development. Against this claim, Carol Gilligan's research—while not definitive—introduces the possibility of a fundamentally

Thus, just as the utopian *social ideal* of critical theory is said to be rendered intelligible in Habermas's theory of discourse ethics because it is rationally based on the normative presuppositions of communicative reason,[68] so the teleological *individual ideal* of psychic maturation is said to rest on a communicatively based tendency toward postconventionality.[69] Habermas writes:

> If by way of a thought experiment we compress the adolescent phase of growth into a single critical instant in which the individual for the first time—yet pervasively and intransigently—assumes a hypothetical attitude toward the normative context of his lifeworld, we can see the nature of the problem *that every person must deal with* in passing from the conventional to the postconventional level of moral judgment. The social world of legitimately regulated interpersonal relations, a world to which one was naively habituated and which was unproblematically accepted, is abruptly deprived of its quasi–natural validity.
>
> If the adolescent cannot and does not want to go back to the traditionalism and unquestioned identity of his past world, he must, on

different, yet equally plausible, moral scheme. The possibility of a different moral voice casts serious doubt on the genuine universality of the Kohlbergian program. See Gilligan's *In a Different Voice: Psychological Theory and Women's Development* (Cambridge, MA: Harvard University Press, 1982).

68. See Albrecht Wellmer's "Reason, Utopia, and the *Dialectic of Enlightenment*," in *Habermas and Modernity*, ed. Richard Bernstein (Cambridge, MA: The MIT Press, 1985), p. 58.

69. This Kohlbergian version of psychic development has remained a consistent feature of discourse ethics. As recently as 1996 Habermas formulated the motivational dimension of his universalist ethic in the following terms. "A principled morality [which] views *everything* through the powerful but narrow lens of universalizability . . . facilitates a knowledge that is meant to orient one's action but does not thereby *dispose* one to act rightly." Significantly, however, he continues, "Sublimated into knowledge, this morality is, like all knowledge, represented at the cultural level. . . . A morality thus withdrawn into the cultural system maintains only a virtual relation to action as long as it is not actualized by the actors *themselves*. . . . A principled morality thus depends upon socialization processes that meet it halfway by engendering *the corresponding agencies of conscience, namely, the correlative superego formations* [my emphasis]. . . . Such a morality becomes effective for action only through the internalization of moral principles in the personality system." Thus, notwithstanding the arguable partiality of the Kohlbergian agent, Habermas is clearly still identifying Kohlberg's version of principled morality as the "normative reference point"—indeed, as the rightful *telos*—of psychic development. See *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, trans. William Rehg (Cambridge, MA: The MIT Press, 1996), p. 113.

penalty of utter disorientation, reconstruct, at the level of basic concepts, the normative orders that his hypothetical gaze has destroyed. . . . Ultimately all that remains is a procedure for a rationally motivated choice among principles that have been recognized in turn as in need of justification.[70]

This shift of attitude "has something unnatural about it," Habermas continues, which, "is like an echo of the developmental catastrophe that historically once devalued the world of traditions and thereby provoked efforts to rebuild it at a higher level."[71] Notwithstanding this unnaturalness, however—or perhaps because of it—the *telos* of sociocultural and ontogenetic development is, in turn, precisely what confirms the link Habermas needs to establish between morality and ethical life. For the consensual end of our argumentative practice, "is not something we can treat so arbitrarily as the contingent ends of action. [It is] . . . intimately interwoven with the intersubjective form of life to which subjects competent in speech and action belong."[72] In other words, Habermas here offers an account in which the strictly formal principle of (U) can be said to link up with the substantive nature of our moral beliefs exactly to the extent that this principle already reflects a substantial interest in psychic integrity that is common to *all* members of communicatively structured societies, and towards which *all* members of such societies must necessarily strive. On Habermas's view morality as such—the rule of impartiality derived from communicative action—addresses precisely the vulnerability which communicative action first creates; since feelings of vulnerability and empathy ensue from an intersubjectively shared web of relations, (U) is said to be normatively binding.[73]

70. "Moral Consciousness and Communicative Action," p. 126; emphasis mine.
71. *Ibid.*, p. 127. See also "Morality and Ethical Life," where Habermas claims, similarly, "Moral universalism is a *historical result*" (p. 208). But compare Habermas's remark that moral intuitions are acquired in a *"quasi–natural manner through socialization,"* in "Discourse Ethics," p. 98; my emphasis.
72. "Discourse Ethics," p. 95.
73. See "Morality and Ethical Life," pp. 202, 203. Elsewhere he writes, "With the validity claims raised in communicative action, an ideal tension is imported into social reality itself, which comes to conscious awareness in participating subjects as a force that explodes the limits of the given context and transcends all

Given the fact that an explicit feature of discourse ethics is its emphatic refutation of value skepticism, this stress on the affective dimension of his moral principle will come as no surprise. The problem with Habermas's argument, however, is that (U) can only be said to confer—not merely *legitimacy* in the legal sense—but what Wellmer identifies as "a corresponding *obligation* to act"[74] insofar as the propensity to universalize is understood in terms of a quasi–natural development. In other words, unless (U) is virtually inescapable, there is no basis for Habermas's claim that the empty procedure of discourse ethics is structurally related to the substantive, ethical interests of all communicative agents. Moreover, unless discourse ethics addresses a "structural feature of the good life" in this sense, Habermas also will not have satisfied the burden of proof regarding his contention that, "my moral principle is not just a reflection of the prejudices of adult, white, well–educated, Western males of today."[75] His response to the charge of ethnocentricism in this case as well thus hinges on the historical and psychological developmental theories he cites, since these are the sole bases of his claim that (U) elucidates a substantive moment of normativity common to *all* forms of sociocultural life.

If Habermas intends to meet the demand for a universally valid normative principle on the basis of a fixed conception of psychosocial development—if this is how he proposes to support his argument concerning the normative force as well as the universal validity of (U)—then it is difficult to see why he perceives a problem of moral justification at all. For in this case there would be no reason to believe that we cannot leave the "gradual embodiment of moral principles in concrete forms of life" to Hegel's absolute spirit after all: there is *in fact* (Kohlberg's analyses are methodologically empirical) no reason to doubt that the "fragmentary realizations" of "the moral intuitions that discourse ethics conceptualizes" will continue to "proliferate" as they already have.[76] In other words, as Seyla Benhabib rightly contends,

---

merely provincial standards." ("Morality, Society, and Ethics," pp. 164–65). Again Habermas's point is that procedural morality bears a normative or moral force which links universal claims of justice to (all) particular notions of the good.

74. "Ethics and Dialogue," p. 158.

75. "Morality and Ethical Life," p. 197.

76. *Ibid.*, p. 208.

"Insofar as the project of communicative ethics is presented as an inevitable sequence of moral development, one reverts back to the philosophy of the subject."[77]

On the other hand, perhaps Habermas does not intend to rely completely on theories of psychosocial development; perhaps Habermas intends for us to overlook his emphasis on the (un)natural tendency toward postconventionality. Certainly there is textual evidence to support this view; Habermas explicitly rejects recourse to "an objective teleology"[78] and, as noted earlier, he is clear that postconventional moral reasoning is far from universally inescapable. In fact Habermas says it requires the support of "socialization processes that meet it halfway by *engendering* the corresponding agencies of conscience."[79] Now if this is so we can well agree that discourse ethics itself is "a moral theory that no longer claims to know the *telos* of "the" good life and which, therefore, "must leave the question 'Why be Moral?' unanswered."[80] As I have argued, however, the logical outcome of this view of discourse ethics is that (U) would specify only "to every participant a numerically equal say"—it would be a rule, in short, which could not account for its own moral (rather than merely logical) status.

On this view, it is true, "postconventional moral consciousness" would not be said to stand alone on the basis of a teleological supposition; rather, it would need to be "supplemented by an enlightened existential self–understanding that entails that I can *respect* myself only as someone who as a general rule performs the actions he takes to be morally right" [*i.e.*, in terms of a common rather than a subjective interest].[81]

77. Benhabib, *Critique, Norm, and Utopia: A Study of the Foundations of Critical Theory* (New York: Columbia University Press, 1986), p. 12. Asher Horowitz makes a more elaborate version of this charge in the course of a discussion of Habermas's use of the concept of performative contradiction. Horowitz writes, "The Subject does not at all disappear within the philosophy of language. The Subject becomes language, or, better, the form of the forms of objectivity given in language." ("'Like a tangled mobile': Reason and reification in the quasi–dialectical theory of Jürgen Habermas," *Philosophy and Social Criticism* 24, no. 1 [1998]: 19).

78. "Morality and Ethical Life," p. 210.

79. *Between Facts and Norms*, p. 113, my emphasis.

80. "Lawrence Kohlberg and Neo–Aristotelianism," p. 127.

81. "Remarks on Discourse Ethics," in p. 80.

Notably, however, two charges can be levelled against this argumentative strategy. First, lacking an inherent *telos*, the theory cannot be said to entail a justified basis for the social imposition of such an enlightenment; the specification of (U) alone does not support the prescription of such a highly circumscribed, collectively–shared understanding of "self–respect." Secondly, to replace a universal *telos* with support from the "existential self–understandings" of the individuals involved would again compromise the autonomy of the moral sphere. For Habermas insists that in "existential–ethical" discourses, "reason and the will condition one another reciprocally, though *the latter remains embedded in the life–historical context thematized.*"[82] Significantly, in other words, if moral consciousness requires existential–ethical supplementation, and if the latter is rooted in the particularities of a given context, then Habermas would seem to be explaining the normative force of the principle (U)—the universally obligatory force underlying the impartial, transcendental point of view—in terms of a *particular* life–historical context. It follows that, without being impelled by the contingencies of one's particular context, there would be no force in "the" moral point of view at all. And this is exactly the position Habermas must reject.

On the view that moral theory makes no teleological claims, moreover, moral consciousness would also need the supplemental implementation of political power for, as Kant already knew, norms are strictly valid only under conditions of their general observance, and "Legal institutionalization alone can ensure general adherence to morally valid norms."[83] Here is it worth citing Habermas's comments in full:

> It is only at the level of a discourse theory of law and politics that we can also expect an answer to the question invited by our analyses: Can we still speak of practical reason in the singular after it has dissolved into three different forms of argumentation under the aspects of the purposive, the good, and the right? . . .The unity of practical reason can no longer be grounded in the unity of moral argumentation in accordance with

82. "On the Pragmatic, the Ethical, and the Moral Employments of Practical Reason," p. 12; my emphasis.
83. "Morality, Society, and Ethics," p. 155; "Remarks on Discourse Ethics," pp. 87–88.

> the Kantian model of unity of transcendental consciousness, for there is no metadiscourse on which we could fall back to justify the choice between different forms of argumentation.
>   ... Moral theory must bequeath this question unanswered to the philosophy of law; *the unity of practical reason can be realized in an unequivocal manner only within a network of public forms of communication and practices in which the conditions of rational collective will formation have taken on a concrete institutional form.*[84]

Here Habermas concurs with Hegel, then, that "unless discourse ethics is undergirded by the thrust of motives and by socially accepted institutions, the moral insights it offers remain ineffective in practice."[85]

Yet such an argumentative strategy is again problematic, for Habermas also concedes that, "in distinction to the moral norms that regulate possible interactions between speaking and acting subjects in general, legal norms refer to the network of interactions in a *specific* society."[86] Indeed, he continues:

> the more concrete the matter at hand, the more the self–understanding of a collectivity and its way of life. . . . are expressed in the acceptability of the way the matter is legally regulated. We see this in the broad spectrum of reasons that enter into the rational process by which the legislature's opinion and will are formed; in addition to moral considerations, pragmatic considerations, and the results of fair negotiations, ethical reasons also enter into deliberations and justifications of legislative decisions.[87]

To this extent, he admits, "every legal community and every democratic process for actualizing basic rights is inevitably permeated by ethics."[88]

The conclusion that can be drawn from these concessions is this: if Habermas is prepared to acknowledge that he cannot appeal to a metadiscourse such as a teleologically oriented theory of ego

---

84. "On the Pragmatic, the Ethical, and the Moral Employments of Practical Reason," pp. 16, 17 (emphasis mine).

85. "Morality and Ethical Life," p. 207.

86. "Struggles for Recognition in the Democratic Constitutional State," in *Multiculturalism*, ed. and intro. Amy Gutmann (Princeton, NJ: Princeton University Press, 1994), p. 124 (emphasis mine).

87. *Ibid.*, p. 125.

88. *Ibid.*, p. 126.

development to justify the choice between different forms of argumentation, and if he acknowledges that his version of "the" moral point of view requires socio–institutional support in order to ensure its proliferation, then there are no grounds on which to privilege moral or right–oriented argumentation as the *foundation* of those same socio–political institutions. In other words, the argument appears to be circular: Habermas is proposing the creation of concrete institutional forms in order to ensure the proliferation of "the" moral outlook. This outlook is said to give rise to universally valid norms. And these norms, in turn, are said to provide the legitimate *grounds* of the very institutions—the political concretizations—which are deemed necessary to ensure the proliferation of "the" moral point of view on which our self–respect is ostensibly based. Clearly, a principled moral outlook of the Kohlbergian sort cannot itself serve—in the name of unconditional validity—as the justificatory grounds of those institutions which are intended to condition or produce it, unless a particular outcome of moral development has been presumed from the start.

Thus Habermas does not actually explain—not, at least, without implicating teleology—how the legitimation principle (U) can be said to constitute the conditions for a so–called dignified human existence.[89] Lacking an *ethical*—which is to say, a culturally and historically–contingent—supposition about moral developmental ends, there is no basis on which to universalize Habermas's understanding of our moral self–respect. Rather, I submit, this particular understanding of morality ethically privileges the figure of a fully autonomous, unambiguously rational, Kantian "Man."

Much more might be said about Habermas's complex and multifaceted program of discourse ethics. This limited discussion of its pivotal features has been intended only to highlight the specific ways in which Habermas relies on the metaphysics of an Enlightenment perspective he intends to have left behind. Thus, just as we have seen that a transcendental notion of the subject is implicated in Habermas's argument for the autonomy of the moral

89. "Morality and Ethical Life," p. 209.

sphere, so I have argued that the teleological dimension of discourse ethics cannot be discarded except at the cost of the theory's very force. Here it emerged that either the normative dimension of discourse ethics is conditioned by a teleological conception of sociocultural and ontogenic development, or else it is without the substantive normative sense that any moral theory is intended to provide. Consequently, Habermas can be said to be no more successful with regard to the task of *re*connecting ethics to morality—the task of refuting the charge of triviality—than he was with regard to establishing the unconditionality of the moral sphere by virtue of their separation.

In both moments of the communicative ethics project, I have argued—in the establishment of an autonomous moral domain as well as in the determination of that domain's normative relevance—Habermas has extricated discourse ethics from the charges of relativism and triviality by multiplying, which is to say, by pluralizing, an Enlightenment concept of the universal subject. The realization of a fully postmetaphysical deontological ethics, therefore, has yet to be accomplished.