Habermas on Moral Justification

Many moral theories have an epistemic orientation: they are preoccupied with uncovering moral truths. This often leads to a quest for moral foundations. By contrast, a practical orientation is attentive first and foremost to the aims of our moral practices—what we are doing when we attempt to establish a morality and to deliberate morally with other people. These aims may include, for instance, the aim of reaching agreement, attending to people’s basic interests, displaying mutual respect, or affirming shared values or a certain way of life. Shifting from an epistemic to a practical orientation can lead us to set aside the quest for the foundations of ethics, since the practical aims of morality could well be pursued and achieved without the support of foundational truths. A practical approach also challenges the divide between “ethics” and “metaethics.” On a practical conception of morality, the moral philosopher is engaged in a task of clarifying and promoting morality’s practical aims. She is a participant, as well as an observer, since the arguments she makes are moral arguments.

Features of what I am calling a practical approach are shared by a range of philosophers that include constructivists such as Rawls and Korsgaard, relativists such as Harman and Rorty, Wittgensteinians such as Putnam and Diamond, and contractualists such as Scanlon and Barry. They are also present in many aspects of Habermas’s “discourse ethics.” Discourse ethics holds that legitimate ethical norms are those that would be affirmed by free and symmetrically situated participants in ideal dis-


© Copyright 2000 by Social Theory and Practice, Vol. 26, No. 2 (Summer 2000)
course. In their deliberations, participants in ideal discourse attempt to determine whether all persons affected by a proposed norm could accept the consequences and side effects that the general observation of the norm could be expected to have for the interests of each person. Habermas elaborates an argument for discourse ethics that aims to establish that the idea of discourse ethics is implicit in the normative regulation of social life, broadly speaking. His analysis of the connection of our social practices with the activity of exchanging reasons in ideal discourse helps us to see what is involved in working out a practical approach to ethics. In exploring these points, this paper thus serves to underscore the interest of Habermas’s thought for some current trends in analytic moral philosophy.

There also seem to be moments in Habermas’s thought, however, in which he is reluctant to give up paradigms associated with the search for foundational moral truths and an affirmation of the “neutrality” of moral philosophy’s observer status. My critical response to Habermas has two parts. The first concerns Habermas’s “justification” of discourse ethics. Habermas maintains that a derivation of the requirements of morality can proceed from the presuppositions of rational communication in general. But I will argue that discourse ethics receives little support from the presuppositions of communication and argumentation per se. Habermas’s account can be seen to depend directly on certain substantive ethical claims. Moreover, these ethical claims could be socially relative; thus, the principles of discourse ethics are not shown to be universally valid, as Habermas’s more ambitious attempt to derive them from the presuppositions of rational communication would imply.

The second part of my critical response (section 3) examines Habermas’s claim that “purely formal” constraints represented by the principles of discourse ethics can nevertheless lead us to particular norms or principles, such as principles of justice. In other words, Habermas thinks that conforming to the principles of discourse ethics will enable us to formulate substantive principles for the regulation of social life. His account of how discourse ethics enables us to converge on substantive regulative principles, however, is not convincing. The “actualization” of a discourse ethics, I contend, relies irreducibly on reasonable judgments about the moral relevance of various sorts of considerations. Discourse ethics must rely on a more robust conception of “the reasonable” than Habermas admits. This points to the ineliminable work of reflection and judgment (a place for philosophy) in a practical approach to morality.

I conclude by defending a strategy for justifying ethics that is less ambitious than the one set forth by Habermas. This strategy nevertheless finds support, I argue, in many aspects of Habermas’s philosophy.
1. The Principles of Discourse Ethics

Discourse ethics claims there is a set of norms that is presupposed by rational communication. The idea is that we commit ourselves to advancing arguments that conform to these norms, if called upon to do so, insofar as we engage in the activity of rationally communicating with other people. When we deliberate with one another about, in particular, the substantive moral norms that may legitimately regulate social life, discourse ethics holds that we are implicitly committed to certain ethical principles (principles U and D, elaborated below). These principles state highly abstract requirements on substantive moral norms and reasons, and should be distinguished from the substantive moral norms and reasons they are meant ideally to regulate. The latter include principles such as principles of justice: a principle that states that all citizens in a liberal democracy are to be granted certain rights and liberties, for instance, must meet the requirements the principles of discourse ethics state, if it is to be considered legitimate. Discourse ethics claims that when we advance arguments that aim for agreement on substantive principles such as these, we have already committed ourselves to the idea that an argument is valid only if it satisfies the ethical constraints articulated by the principles of discourse ethics. I will now spell out these claims in some detail by representing Habermas's argument in six steps:

1. Rational communication is a form of communicative action. Habermas claims that rational communication has a practical purpose: it serves to coordinate the “action plans” of several actors. It does so, however, in a particular way. Communicative action takes place “when actors are prepared to harmonize their plans of action through internal means, committing themselves to pursuing their goals only on the condition of an agreement—one that already exists or one to be negotiated—about definitions of the situation and prospective outcomes.” By referring to “internal means,” Habermas intends to exclude the use of force or sanctions to bring about agreement in judgment and harmony of action plans. Communicative action relies instead on the free consent of the participants. More precisely, communicative action aims to reach a mutually acceptable agreement that will serve to render harmonious the action plans of the actors involved.

---

2^The justification of substantive moral norms accepts the principles U and D of discourse ethics as basic; hence it will not be circular (but would be dogmatic, if the principles of discourse ethics cannot themselves be justified).


4^Ibid.
Communicative action contrasts with strategic action. This contrast concerns the motivational impact and methods of acting. Strategic action is “action oriented toward success” given egocentric aims. Strategic rationality involves instrumental calculations: the problem is to achieve some end by selecting the appropriate means. Strategic action may at the same time be social (rather than purely technical) in that it can involve an agent’s attempt to influence the behavior of other people, but strategic attempts to influence others lack the value placed on reaching consensus that is characteristic of communicative action. The promise of rewards and outright threats of harm as well as other forms of coercion implemented for strategic purposes are not necessarily irrational, and consequently social relations characterized by strategically rational actions are typically conflictual and unstable. The coordination of actions and any degree of cooperation depend solely on the extent to which the self-interested utility calculations of the actors involved mesh.

Communicative actions, on the other hand, are actions in which agents are oriented toward reaching a mutual understanding, and each appeals to a like orientation in others. Habermas writes:

Whereas in strategic action one actor seeks to influence the behavior of another by means of the threat of sanctions or the prospect of gratification in order to cause the interaction to continue as the first actor desires, in communicative action one actor seeks rationally to motivate another by relying on the illocutionary binding/bonding effect (Bindungs-effekt) of the offer contained in his speech act.

The “illocutionary binding/bonding effect” Habermas refers to is the motivational impact on others of the promise each participant implicitly makes to provide reasons sufficient to defend the validity of his or her claims, if challenged to do so. The effect is that, in general, agents assume that the claims others make are valid. This leads to our second step.

2. Communicative action involves the aim that others accept the validity of our utterances. As indicated above, communicative action is action “oriented toward reaching understanding.” When we speak or otherwise act with the intention that another person understand us, Habermas’s “formal pragmatics” claims that we are attempting to do

---

5On Habermas’s contrast between strategic and communicative action, see Moral Consciousness and Communicative Action, pp. 58, 133-34.

6In The Theory of Communicative Action, Vol. 1 (trans. Thomas McCarthy; Boston: Beacon Press, 1974), Habermas draws a contrast between strategic and instrumental action (see p. 285). I will not be concerned with the subtleties of the distinction drawn there.

7Habermas, Moral Consciousness and Communicative Action, p. 133.


9Habermas, Moral Consciousness and Communicative Action, p. 58.

something. We are aiming to establish a relationship with that person that depends on his or her acceptance of the validity of our utterance. The consensual regulation of action achieved through communicative action supposes, Habermas thinks, that the consensus reached (or aimed at) could at any point be evaluated "in terms of the intersubjective recognition of validity claims." Validity claims have cognitive content, he argues, that is given in the form of reasons. Thus consensus on scientific truths or moral norms, for example, can be evaluated in terms of the reasons agents take to support them.

3. A claim p is valid iff p would be accepted through rational argumentation or discourse. Rational actors can at any time challenge each other to give reasons for the validity of their claims upon one another. Thus, communicative action always points toward the possibility of rational argumentation. Habermas writes:

The rationality proper to the communicative practice of everyday life points to the practice of argumentation as a court of appeal that makes it possible to continue communicative action with other means when disagreements can no longer be repaired with everyday routines and yet are not to be settled by the direct or strategic use of force. When such disagreements arise, the participants can "pass over to the level of discourse." In discourse, reasons are put forth and defended with arguments. The strength of an argument is evaluated in terms of the soundness of the reasons given. Habermas claims that like communicative action, discourse is "domination-free." In discourse (unlike in communicative action in general), what this means is that nothing but the force of the better argument motivates its participants to accept a particular outcome.

As suggested earlier, in communicative action that has not yet passed over into the level of discourse, the motivational impact of communicative action is not directly provided by the actual cognitive validity of a speaker’s claims but rather is sparked by a "speaker’s guarantee that he will, if necessary, make efforts to redeem the claim that the hearer has accepted." The speaker affirms the validity of his or her claims and expresses a commitment to defending that validity claim with sufficient reasons, if called upon to do so. Each judges and acts only in ways that he or she believes can be justified to others (that is, all claim to be able "discursively to redeem" their claims to validity). The effect is that, in

---

11Habermas, Moral Consciousness and Communicative Action, p. 58.
15Habermas, Moral Consciousness and Communicative Action, p. 58.
general, agents assume that the claims others make are valid. The hearer’s rational acceptance of the speaker’s guarantee enables consensus of judgment to come about and the rational coordination of action plans to be possible.

4. In practical discourse (vs. discourse in general) we assess the validity of proposed norms in accordance with a Discourse Principle (D). Habermas calls the context for the evaluation of reasons with respect to action-guiding norms practical discourse. Members of modern, pluralistic societies, Habermas writes, “find themselves embroiled in global and domestic practical conflicts in need of regulation.” They continue to regard these conflicts as rationally resolvable, “but their shared ethos has disintegrated.” For a possible resolution, discourse ethics abstracts from the content of people’s disagreements, and directs them to a consideration of the formal or structural features of discursive practices per se. Habermas claims this leads us to a “Discourse Principle” (Principle D). Principle D states: “Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity as participants in a practical discourse.” D articulates, in a purely formal way, an understanding of what it is to justify action-guiding norms. It conveys the basic idea of discourse ethics.

The introduction of Principle D is, however, hypothetical in the sense that the possibility of successfully instantiating this principle needs to be established. Formidable obstacles to its successful application are found in the very conflicts that give rise to the need for discourse ethics in the first place: familiar and sometimes deep disagreements among persons with different value-orientations about how to resolve normative questions. If discourse ethics is to justify its claim that the formal presuppositions of discourse can be used to resolve such disagreements, it must es-

---

18Ibid. This is also a central theme of Rawls’s Political Liberalism.
19Ibid., pp. 40-41.
20Habermas, Moral Consciousness and Communicative Action, pp. 66, 93. It is not always clear whether Principle D is itself supposed to be derived from an examination of the formal features of discourse (See Jürgen Habermas, Justification and Application: Remarks on Discourse Ethics, trans. Ciaran P. Cronin (Cambridge, Mass.: The MIT Press, 1993), pp. 49-50), or whether it is presented as a hypothesis that is later confirmed by exploring the implications of the formal features of discourse (See Habermas, Inclusion of the Other, pp. 41-43). Here I will assume the latter.
21I am here following Habermas’s presentation in Inclusion of the Other, where Habermas says that Principle D expresses a “conception of normative justification in general” (p. 43). In that work, he formulates D immediately after proposing that “the appeal to moral content must be replaced by the self-referential appeal to the form of [discursive] practice” (p. 41).
Habermas on Moral Justification

Establish a principle that can show how conflicts are to be negotiated. In order to "operationalize" $D$ with respect to moral discourse, that is, to show that it can be realized, Habermas claims a "bridging principle" must be introduced that can establish the possibility that agreement on action-guiding norms could in fact be reached in the face of value pluralism. The derivation of this bridging principle will thus respond to the condition of pluralism characteristic of modern (liberal) societies. Habermas's approach is the following: he aims to establish that the derivation of this bridging principle depends on claims about the formal features of argumentation in general.

5. The formal features of rational argumentation in general are captured by "Alexy's rules." As we have seen, Habermas believes that communicative action points toward the possibility that the validity claims implicit in such action can be redeemed through argumentation. This is true, he believes, of communicative action of all sorts. As he understands it, the argumentation toward which communicative action points is not exclusively moral argumentation about the validity of norms for regulating actions; it may be such as to address truth claims more broadly construed. As Habermas understands it, communicative action encompasses all illocutionary acts. He analyzes illocutionary acts as falling into three general categories. First of all, they encompass constative speech acts, which are speech acts that put forth claims to truth (cognitive claims). Secondly, they encompass subjectively expressive speech acts, which can be evaluated for their authenticity. And finally, they encompass regulative speech acts, which aim to establish the validity of norms to govern social interactions. Illocutionary acts, as Habermas understands them, include all acts we perform with language that aim to establish a certain sort of interpersonal relationship: one in which the hearer understands the claims the speaker puts forward, and accepts them as valid. All such acts could in principle call upon speakers to back up their claims with arguments.

Drawing upon a general theory of argumentation, Habermas maintains that all argumentation presupposes that a context of discussion guarantees in principle freedom of access, equal rights to participate, truthfulness on the part of participants, and the absence of coercion in adopting positions. He refers to these as "general symmetry conditions." Following Robert Alexy, he states these presuppositions more formally as follows:

---

22They also include what Habermas calls claims to "truthfulness," that is, claims about the authenticity of conceptions of value and the good life.

23Habermas, *Justification and Application*, p. 31.
1. Every subject with the competence to speak and act is allowed to take part in a discourse.
   a. Everyone is allowed to question any assertion whatever.
   b. Everyone is allowed to introduce any assertion whatever into the discourse.
   c. Everyone is allowed to express his attitudes, desires, and needs.
2. No speaker may be prevented, by internal or external coercion, from exercising his rights as laid down in 1 and 2.24

I shall refer to these presuppositions as “Alexy’s rules.” The theory of argumentation upon which discourse ethics relies alleges that arguments that depend on violating any of these rules fall short of being rationally convincing, and thus must fail to demonstrate the validity of principles or other claims put forth. If the above presuppositions are in fact unavoidable and rationally necessary premises of argumentation, as Habermas claims, then any attempts to jettison them will be inconsistent. Consider the following statement: “By excluding questions that slowed down the discussion, we demonstrated that P is rationally defensible.” Such a statement, argues Habermas, involves a “performative contradiction.” Argumentation that claims to establish the validity of a claim implies that a mutual agreement on that claim by way of the full discursive participation of rational and relevantly situated persons could be worked out. In this example, however, it is claimed that validity has been established under conditions that preclude the free and full participation of all. This is inconsistent.

Alexy’s rules are idealizing conditions. The idea is that all persons, insofar as they exercise their capacity as participants in discourse, have certain rights. Habermas’s claims about the nature of practical discourse are, however, consistent with the fact that many of our interactions with one another contain some elements of strategic as well as communicative action. A speaker’s intentions must be communicative if her action is to succeed in bringing about an interpersonal relation based on consensus, but this need not be true of all persons involved. The presuppositions of communicative action are then suitably modified. Consider the following statement: “Although we excluded dissenters from the discussion, we managed to demonstrate the validity of the reasons under consideration.” Suppose we interpret the speaker to mean that ideally all persons motivated to reach consensus through domination-free discourse could have agreed upon the validity of the claims in question, but in fact persons with contrary strategic aims interfered with the possibility of such discourse, and were therefore excluded. Discourse ethics could hold that

one might plausibly be able to defend the validity of a claim in this way. Argumentation that proceeds by insisting there could be no good objections that are not demonstrably strategic in nature, however, would not be rationally convincing, and could not establish the validity of the claims in question.

6. Finally, from Alexy’s rules and some idea of what it is to justify norms (D), we can deduce Principle U. Habermas holds that Alexy’s rules, together with a weakly normative understanding of what it means to justify norms,25 entails a “Principle of Universalization” (Principle U). His argument is that if one engages in argumentation with the aim of establishing that certain norms of action ought to be adopted, that is, if one attempts to redeem the validity claim associated with regulative speech acts, the presuppositions of discourse (captured by Alexy’s rules) entail that one also presupposes Principle U. Principle U claims: “A norm is valid when the foreseeable consequences and side effects of its general observance for the interests and value-orientations of each individual could be jointly accepted by all concerned without coercion.”26 This principle, Habermas claims, part of the logic of moral argumentation.27

What U appears to add that D lacks is a more specific analysis of the moral status of persons and their claims: the interests of each person are morally important and can ground moral claims upon us. All persons have equal claims, based upon their fundamental interests. U acknowledges this equal status by having us take up the perspective of other persons and determine whether the proposed norms are acceptable in view of the interests of all. In effect, then, it directs each person to take an interest in the interests of others. This reciprocal perspective-taking establishes that all can accept proposed norms when each considers the effect that general compliance with those norms would have for the interests of all. Moral norms, thus, are not the outcome of self-interested bargaining and compromise, but represent a genuine consensus that each participant considers fair to and in the best interests of all.

25As claimed in n. 21, Habermas’s recent work suggests that this weak understanding of what it means to justify norms is expressed in Principle D. Previously, he formulates this understanding more loosely as the idea that “moral justifications resolve disputes concerning rights and duties, that is, concerning the rightness of the corresponding normative statements,” and that the resolution of such disputes takes place via rational deliberation (i.e., on the basis of reasons). Justification and Application, p. 32. See also p. 52, and Moral Consciousness and Communicative Action, pp. 92-93.

26Habermas, Inclusion of the Other, p. 42. For alternative statements of the principle, see Habermas, Moral Consciousness and Communicative Action, pp. 65, 93.

27Habermas, Moral Consciousness and Communicative Action, p. 57.
2. The Justification of Discourse Ethics

In this section of the paper I examine the role that the theory of argumentation plays in the derivation of Principle $U$. Habermas maintains that the theory of argumentation provides a broad base for Principle $U$. I question this.

Alexy's rules, as I understand them, are meant to spell out reasonable constraints on arguments suited to discourse in general, constraints that must be met by all valid arguments. But it is implausible to suppose that there can be a single, unified account of criteria for specifying when it is that agents are reasonably situated to evaluate validity claims, an account that will be fitting with regard to all discourses. Why should we think that theoretical discourses, for instance, must satisfy these rules? Alexy's rules imply that persons are reasonably situated when they have the right fully and freely to participate, introducing whichever considerations they see fit. It is far from clear, however, that argumentation for the validity of truth claims in science, mathematics, and logic, for instance, aims to be communicative in a sense that would impose these requirements. We do not think that anyone ought ideally to be able to introduce any considerations whatever into rational discourse on these subjects. The truths with which theoretical discourse is concerned need not be truths to which anyone could consent. While greater inclusion may well be to the benefit of science, this would not establish that it is a criterion of truth that truth claims could be the object of an all-inclusive consensus. Although Habermas does not ignore these matters, it is not clear that he provides a convincingly detailed account of how Alexy's rules apply to theoretical discourse.

Given the surface implausibility of the application of Alexy's rules to...
theoretical discourse, we should ask, why is it important to Habermas's argument for discourse ethics that the theory of argumentation he relies upon in deriving $U$ applies to all discourse? Habermas is concerned to establish that the idea of discourse from which he claims to derive presuppositions that lead us to Principle $U$ be broad and not ethically framed. He writes: "The significance of these [formal pragmatic] arguments is proportional to the degree of generality of the discourses that entail substantive normative presuppositions." The idea seems to be this: the more general the premises, the more the derivation of normative presuppositions establishes. This idea is connected with Habermas's aim to avoid an objection he levels at some others. He aims to avoid the objection that in the presuppositions of discourse the theorist finds only the normative substance that he had previously put into an implicit definition of practical discourse. He writes: "The delineation of an object domain must not already prejudge the normative content of its presuppositions, or one will be guilty of a *petitio principii* that could have been avoided."

These remarks indicate that Habermas is not simply attempting to show that our commitment to the principles of discourse ethics is unavoidable, insofar as we engage in practical discourse. He wants, beyond that, to establish that the normative content of the principles of discourse ethics itself has a *justification*. Paradoxical as it may seem, he aims to show that the fact that argumentation advanced in practical discourse presupposes the validity of discourse ethical principles can also be used to account for the *authority* such argumentation has to provide ethical justifications. In other words, he wants to give an account of how it is that the principles of discourse ethics have the authority to justify norms.

Discourse ethics is intended as an alternative to a foundationalist picture of moral justification. On a *foundationalist* picture, justification is understood as linear and ultimately non-inferential. The foundationalist claims that through reason and reflection we arrive at the conclusion that certain principles are fundamental and need no further justification, and that they can be used to justify other, less foundational, principles and reasons. The burden of justification is carried by first principles and other abstract (highly general) conceptions alone. Proponents of discourse

---

31Ibid., p. 84.
32Ibid., p. 83.
33Habermas ultimately rejects the possibility of a "final justification," but I shall argue that he does so with ambivalence.
35This concerns what Rawls refers to as the "order of support" for a moral concep-
ethics argue that foundationalism is misguided.\textsuperscript{36} Alexy, for example, argues that the demand for ever more fundamental justifying claims (ever more basic substantive principles) is replaced in discourse ethics by a rational demand on the activity of justification itself.\textsuperscript{37} What this means is that principles that justify particular judgments and actions are principles that must themselves meet certain demands or requirements of reason.\textsuperscript{38} The rational demand discourse ethics places on the activity of justification is that an agent's reasons satisfy the ethical principles presupposed by the very possibility of practical discourse. Attempts to evaluate the discourse ethical principles themselves belong as well to practical discourse; thus, any attempts to argue in favor of (or against, for that matter) the validity of the (meta-)ethical principles discourse ethics identifies will already presuppose the validity of those principles.

Habermas is worried, however, that the “justification” of Principle \textit{U} avoids foundationalism only to encounter a \textit{petitio principii}. This worry seems to make him reluctant to relinquish altogether the idea that justification has a foundationalist structure. As suggested above, in order to avoid being charged with a \textit{petitio principii}, he aims to derive discourse ethical principles from premises that are both unavoidable (because highly general) and do not themselves have determinate ethical content (or at least do not have ethical content we can “prejudge”), but this move takes him back towards foundationalism. A tension emerges within Habermas’s justification of discourse ethics: although his “postmetaphysical” philosophy attempts to repudiate foundationalism, his “formal pragmatics” appears to take on the role of foundationalist “first philosophy.”

The considerations that support \textit{U}, however, reveal that \textit{U} is not rooted in rational communication generally, as Habermas claims it is. As discussed above, Principle \textit{U} claims that a contested norm cannot meet with the consent of the participants in a practical discourse, “unless all affected can freely accept the consequences and the side effects that the \textit{general} observance of a controversial norm can be expected to have for the satisfaction of the interests of each individual.”\textsuperscript{39} Implicit in Principle

\textsuperscript{36}Their charge is that foundationalism is dogmatic. See Karl-Otto Apel, “Is the Ethics of the Ideal Communication Community a Utopia? On the Relationship between Ethics, Utopia, and the Critique of Utopia,” in Benhabib and Dallmayr (eds.), \textit{The Communicative Ethics Controversy}, p. 42.


\textsuperscript{38}There is a close connection here with Kant’s argument that the only way to avoid an infinite regress is to turn from an examination of the \textit{content} of an agent’s maxim to evaluate instead its \textit{form}. See Immanuel Kant, \textit{Groundwork of the Metaphysics of Morals}.

\textsuperscript{39}Habermas, \textit{Moral Consciousness and Communicative Action}, p. 93.
$U$ is an idea of the *reasonable* conditions for consent: each agent has interests that any consensus must respect, if the agent's consent is to be reasonable. This claim is rather loosely formulated; Habermas does not specify the nature of people's interests, leaving that to the task of practical discourse. What he does suggest is that it is reasonable for an agent to consent when the agent can consent freely (i.e., without coercion) in view of her understanding of what her interests are, and while understanding what the consequences of general compliance with the proposed norm would be for these interests.

However vaguely specified it is, a normative claim that was not implicit in the prior premises has been added with the introduction of Principle $U$: the claim that each person must be able to accept the consequences that the acceptance of proposed norms would have for his or her interests.\footnote{Here and also below I disagree with some critics, such as Benhabib, who believe that $U$ adds nothing substantive to Alexy's rules. See n. 50.} If Alexy's rules are meant to codify the presuppositions of rational argumentation in general (and I have questioned whether they succeed in doing so), the condition of rational consent based on interest cannot plausibly be among them. The validity of logical inference, for example, or the truth of the laws of nature, does not depend upon whether it is compatible with, or supported by, the interests that participants in discourse might affirm.\footnote{Rehg is similarly skeptical about whether the idea that participants should be allowed to express their needs and desires can be considered a general rule of argumentation. See his *Insight and Solidarity*, p. 64.}

Admittedly, Habermas does not claim to be deriving Principle $U$ from the presuppositions of communicative action alone. He claims to be deriving them from the presuppositions of communicative action *plus* an understanding of what it means to justify norms (weakly construed). But it appears that it is further philosophical analysis of what is involved in justifying norms to others that gives Principle $U$ its normative content: the role Principle $U$ admits for the idea of "the reasonable" is characteristic of discourse that is *morally motivated*. Principle $U$, as I am interpreting it, expresses the idea that the ideal (practical) discourse situation is restricted to persons who are motivated to justify their actions on grounds others could freely and rationally accept in view of their interests, and this means that only certain reasons should be advanced within the discourse.\footnote{The similarity between this idea and Scanlon's contractualism will become clear in the text below.} In particular, participants should advance in support of proposed norms only reasons that respect the interests of all. It is this substantive moral claim that potentially provides a "bridge" to agreement, but it is hard to see how it represents a formal feature of discourse
in general. The derivation of $U$ does not appear to be grounded in rational communication as such.

3. The Possibility of Consensus

The idea of communicative action that is directed toward the moral end of establishing harmonious social life by way of achieving consensus on regulative norms is, I have stressed, narrower than the one Habermas puts forth as crucial to his argument. If we interpret him as analyzing this more restricted sense of communicative action—one associated with the requirements of moral reasoning, in particular—his view would be much closer to a view like T.M. Scanlon's. Scanlon's contractualism presents an account of certain reasonable conditions on moral discourse as one that distinguishes social interaction that is morally motivated. Moral motivation, Scanlon argues, is motivation to justify one's actions to others on the basis of reasons they could not reasonably reject. The idea of reasonable rejection he develops has explicit moral content.

Followers of Habermas allege that by anticipating the normative content of moral discourse, Scanlon's contractualism is not sufficiently "dialogical" and that it fails to require that discourse be "actualized." I will now investigate what the "actualization" (i.e., implementation) of ideal discourse involves, and whether this is something Scanlon ignores. Along the way, I will develop some objections to the practical relevance of Habermas's discourse ethics. Although Habermas stresses that discourse is open-ended—that it is, in theory, potentially never-ending—he also believes the principles of discourse ethics can often in fact guide deliberation to a conclusion. Specifically, he claims that when Principle

---


45Habermas does not foresee difficulty in generating consensus on norms (vs. their
U is adequately contextualized, it can be used to distinguish valid norms. Thus, the actualization of practical discourse is linked to securing concrete results, that is, agreement on particular action-guiding norms. I doubt that the "purely procedural" constraints of discourse ethics provide sufficient guidance to enable this. Discourse ethics must instead rely upon a notion of relevance.

References to free and unfettered discourse suggest that conforming to Alexy's rules itself brings about agreement on the validity of proposed norms (or at least promotes it). This idea would presumably be endorsed by those interpreters, such as Benhabib, who think that Principle U adds nothing of substance to Alexy's rules. Habermas himself writes that "only [in] an actually carried out discourse... is the exchange of roles of each with every other forced upon us." Recall first the ideal conditions of any discourse situation Alexy's rules purport to articulate. From the nature of the aim of communication implicit in illocutionary acts generally, Alexy and Habermas claim to derive, among other constraints, "general symmetry conditions" on ideal discourse. Discourse ethics (at least on this interpretation) then appears to maintain that consensus legitimately attained in discourse is any consensus that would (or could) be reached if anyone were allowed to participate in discourse and could advance any claims whatever. The ideally symmetrical situation of participants in practical discourse ensures that all are allowed to take part in the discourse and equally to express themselves within it. Further-


Habermas claims that practical discourses "depend on content brought to them from outside." See Moral Consciousness and Communicative Action, p. 103. On the importance of a "background consensus," see Between Facts and Norms, p. 22.

I thank Joshua Cohen for helping me to formulate this point.

See, for example, Chambers, Reasonable Democracy, pp. 168-72; McCarthy, The Critical Theory of Jürgen Habermas, p. 326.

Benhabib writes: "[S]ince whether or not 'the content and field of application of a norm is or is not particular' can only be discovered in argumentation, it is unclear what the universalizability principle adds to the argumentation procedure." Benhabib, Critique, Norm and Utopia, pp. 303-4. Benhabib also doubts that Alexy's rules capture the presuppositions of argumentative speech generally, and argues that they have cultural-historical content built into them (p. 306). See also her "Liberal Dialogue Versus a Critical Theory of Discursive Legitimation," in Nancy L. Rosenblum (ed.), Liberalism and the Moral Life (Cambridge, Mass.: Harvard University Press, 1989), pp. 143-56. Although Benhabib appears to favor the interpretation according to which U adds nothing to Alexy's rules, she also considers the possibility that it adds some substantive moral content. She concludes that "the presence of a universalizability principle in communicative ethics is either redundant or inconsistent." See Critique, Norm and Utopia, p. 308.

more, none will be forced to accept proposed principles; the conclusions reached can thus be said to have been freely accepted by all.

Earlier I questioned whether these symmetry conditions are required by theoretical discourse. I am now raising a different objection: compliance with Alexy’s rules does not ensure that it will be possible to reach an agreement under those conditions (even in the more restricted case of practical discourse), much less that any agreement reached would be reasonable. If anyone can participate, and can introduce any considerations whatsoever, there is little reason to believe that the parties can be expected to agree. The idea that the application of Alexy’s rules alone can be expected to generate regulative norms is not convincing.

I have argued that Principle U conveys substantive moral content not contained in Alexy’s rules, and that it is derived on the basis of premises not found in Alexy’s rules. Unlike Alexy’s rules, Principle U formulates an idea of the aims it is reasonable for participants in (ideal) discourse to attempt to advance: participants should propose as valid only principles they believe express equal respect for and consideration of the interests and value-orientations of all. Thus, if participants do not conform to this requirement, they violate the conditions of ideal discourse. This narrows the range of considerations participants in (ideal) discourse can put forth, increasing the likelihood of convergence. In this way, U articulates substantive normative constraints on proposals that are worth taking seriously. That is, what count as legitimate reasons within discourse must be compatible with a certain ideal of equality. But then it would appear no longer to be true that in ideal discourse, “anyone can advance any claims whatever.” Certain sorts of proposals will be discounted, even though they may generally be compatible with the communicative aim of persuading others on the basis of reasons. Such proposals could be those that fail to draw upon shared values and interests, appealing instead, for instance, to the aggregate good, to the value of particular aesthetic or religious conceptions, or to the importance to the majority of an established way of life. They could be proposals that fail to rely on empirical evidence (when relevant) and may depend objectionably on private experience or anecdotal evidence. Or they could be dogmatic about how to understand key concepts, the priority ranking of shared values, the appropriate weight of relevant considerations, or the ways in which experience supports value judgments.

---

52 As Joshua Cohen argues, constraints on reasons that can be advanced in a fair procedure limit the substantive outcomes of the process; procedural and substantive values cannot be separated. See his “Pluralism and Proceduralism,” Chicago-Kant Law Review 69 (1994): 609-10.

53 See Rawls’s discussion of what is required in order to recognize the “burdens of judgment” in Political Liberalism, pp. 54-58.
The point is that once we understand how the substantive moral content of Principle $U$ extends beyond Alexy's rules, we are in a position to see that Alexy's rules do not provide us with any real guidance in determining which claims we have moral reasons to take seriously. Principle $U$ tells us that the claims people advance in practical discourse need to be assessed for their moral content in order to determine their relevance: we need to assess whether proposals express consideration of the interests and value-orientations of all and thus could be affirmed in a procedure of universal perspective-taking. The mere fact that somebody has freely advanced a claim in discourse carries no weight. There would even appear to be some tension between promoting Alexy's rules and the discriminatory judgment required by $U$.

An advantage of what we might think of as the "formalist" interpretation of $U$—which more or less equates the content of $U$ with the alleged formal features of discourse elaborated by Alexy's rules—is that it makes it easier to envisage what would be involved in actualizing practical discourse procedurally governed by $U$. But read as a substantive moral principle, $U$ cannot be implemented as straightforwardly as Alexy's rules can be. The actualization of $U$ requires the subtle and discriminatory judgment required in perspective-taking and the conduct of fair play. The moral sensitivities of those persons capable of making use of $U$ as a regulative principle must already be considerably developed.\(^{54}\) Participants in discourse must understand the difference between what is respectful and what is inconsiderate, exploitative, or humiliating, for instance, if they are to succeed in implementing $U$.\(^{55}\) They must maintain a sense of fairness and good will toward one another, and be capable of discerning which proposals fall short of demonstrating full commitment to the aims of practical discourse. Also, as we have seen, they must be able to make out clear and cogent arguments that are responsive to the relevant information and careful in addressing matters that are commonly the subject of disagreement.

Let us return briefly to a comparison between Habermas's discourse ethics and Scanlon's contractualism. Scanlon claims that a contractualist notion of moral wrongness can be stated as follows: "An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one

---

\(^{54}\)For an interesting discussion of the moral sensitivity required to see the relevance of moral principle in concrete life situations and to frame maxims adequately for evaluation by moral principle, see Barbara Herman, "The Practice of Moral Judgment," in *The Practice of Moral Judgment* (Cambridge, Mass.: Harvard University Press, 1993), pp. 73-93.

\(^{55}\)See Hilary Putnam, "Values and Norms," unpublished manuscript, as well as Herman, "The Practice of Moral Judgment."
could reasonably reject as a basis for informed, unforced general agreement." On a contractualist view of moral wrongness, an action fails to be justifiable to others if it is disallowed by a system of rules which no one could reasonably reject provided that they are also motivated to find such principles. An action is justifiable to others, on the other hand, if it is allowed by a system of such rules. Like the principles of discourse ethics, the contractualist formula is a principle to guide the choice and evaluation of first-order normative principles and reasons. The formula holds that we may say that a person's reasons are justifiable to others provided that those others cannot reasonably reject principles which allow her to act in accordance with those reasons—principles that could be the object of informed, unforced agreement between the concerned parties. Determining whether a principle can reasonably be rejected, in contractualist terms, requires us to engage in the sort of perspective-taking Habermas identifies with Principle U. The contractualist formula is more or less equivalent in content to Principle U.

In analyzing the requirements of contractualism, Scanlon stresses the sense in which moral discourse is always guided by a substantive idea of "the reasonable." Moral discourse proceeds by identifying not which proposals are actually rejected, but which proposals could reasonably be rejected. This does not mean that the results of discourse can be anticipated a priori. The philosopher tries to anticipate, given appropriate background information about a given "lifeworld" (to use Habermasian terminology), what reasonable and morally motivated persons would actually reject. This appears to Habermas to be unacceptably monological. But as participants in discourse, this is how we must proceed in evaluating actual proposals. The actualization of the moral principle expressed by Principle U and the contractualist formula cannot avoid making use of this sort of abstract reasoning. How else could we evaluate one another's claims?

Scanlon's contractualism holds that the significance of basic interests for moral discourse is that anyone could reasonably reject principles that neglect his or her basic interests. This idea is similar to Habermas's idea that it would not be rational for persons to agree to principles that do not take their interests into account. Moreover, both would agree that which

---

57Scanlon's recent book argues that the idea of what cannot reasonably be rejected could be replaced by the idea of what could reasonably be accepted without altering the content of the contractualist doctrine. See What We Owe To Each Other, chap. 3.
58Rawls's theory of justice can be understood along these lines. See A Theory of Justice (Cambridge, Mass.: Harvard University Press, 1971), and Political Liberalism. See also Barry's discussion of the importance of combining an a priori and an empirical approach in Justice as Impartiality, chap. 8.
interests are morally basic is something that must be determined within moral discourse. In Habermas's terminology, this involves determining which interests are "generalizable": interests whose legitimacy can be affirmed from all points of view. Scanlon claims that in moral reasoning we should appeal only to "generic" reasons. Generic reasons are "reasons that we can see that people have in virtue of their situation, characterized in general terms, and such things as their aims and capabilities and the conditions in which they are placed."59 Such reasons, like Habermas's generalizable interests, cannot be attributed only to specific individuals, picked out by proper names or "rigged" definite descriptions; the description of relevant reasons/interests must fit into more general categories the moral legitimacy of which could be the object of a broad consensus. Scanlon claims that generic reasons include such reasons as "to avoid bodily injury, to be able to rely on assurances [we] are given, and to have control over what happens to [our] own bodies."60 There is, admittedly, a realist edge in his reference to the generic reasons we have (we "see" them). In an earlier paper, however, he develops a discussion of "urgent" needs as generally shared needs whose content can be constructed in moral argumentation (via reaching agreement).61 That conception of needs is clearly compatible with contractualism. In any case, the realization of both Principle U and the contractualist formula relies on the exercise of comparative judgments about the moral importance or unimportance of various professed interests.

Even after we arrive at a conception of generalizable or generic interests, contractualism holds that it is often crucial to moral reasoning that it involve certain kinds of further comparisons. We need to determine what it is reasonable to ask of others, even in view of one's morally legitimate interests, in order to account for the possibility of consensus. Scanlon argues, for instance, that comparative judgments about the relative severity of deprivations that would result from adopting a proposed principle (as opposed to its alternatives) could provide reasonable grounds for rejecting it.

Contractualist morality relies on notions of what it would be reasonable to accept, or reasonable to reject, which are essentially comparative. Whether it would be unreasonable for me to reject a certain principle, given the aim of finding principles which no one with this aim could reasonably reject, depends not only on how much actions allowed by that principle might hurt me in absolute terms but also on how that potential loss compares with other potential losses to others under this principle and alternatives to it.62

59 Scanlon, What We Owe To Each Other, p. 204.
60 Ibid.
Morality often requires us to make judgments about the relative urgency of goods for particular persons. Generally speaking, more urgent needs should take priority. For example, it may well be unreasonable (as well as irrational) for a person to accept severe deprivations for herself in order to confer a much smaller benefit on each member of a larger group, even if that person would freely accept the sacrifice. Accepting a person’s willingness to make such a sacrifice could (although it may not always) involve taking advantage of that person. Moral judgment is required in order to “actualize” Principle U, that is, the actualization of U relies (irreducibly) on comparative judgments of the costs it is reasonable to impose on some persons for the sake of benefits to others.

Habermas is right to maintain that there are often many choices that participants in practical discourse can legitimately make (that is, as he sees it, there are many choices that would not generate performative contradictions). He also seems right to claim that a determinate result can be reached only when the various individual and shared purposes and background understandings are figured in, that is, when discourse (approximating ideal conditions) is carried out. The question is how these purposes and understandings are to figure in and how it is reasonable to proceed in view of them. This is where the real work of moral reasoning is done. Habermas takes this to show that principles best suited to resolve a practical moral conflict must be determined by what (and whatever) the particular persons who are engaged in the conflict would actually choose were they to abide by the constraints of ideal discourse. This is what he means when he says that discourse ethics is purely procedural and that the principles of discourse ethics require contextualization. But now we can see that what it is to abide by the constraints of ideal discourse is far from straightforward, even if we agree that affected persons should have the last word on what counts as a reasonable set of norms for them.

The possibility of moral consensus depends on people’s commitment to finding reasonable and even creative ways to balance their competing yet legitimate interests. The notion of the reasonable (thematized by Scanlon and others) plays a larger role in actualizing discourse ethics than Habermas would seem to admit. It characterizes not only the

---

63See Scanlon, “Preference and Urgency.”
64Scanlon expresses agreement with this point. See What We Owe To Each Other, p. 206. Herman argues that contextualization of this sort is required in order to apply Kant’s Categorical Imperative as well. See “The Practice of Moral Judgment,” especially pp. 89-90.
66On the role of the reasonable in moral discourse, see also Charles Larmore, The
symmetrical situation of participants in moral discourse, but also their orientation and aims within it. It is because Scanlon is involved in thinking about how to actualize ideal discourse that his contractualism thematizes "the reasonable." Principle $U$, like the contractualist formula, in effect holds that the parties to practical discourse should aim not for agreement per se, but for reasonable agreement. The fact that an agreement emerges from discourse characterized by Alexy's symmetry conditions is not sufficient to ensure its reasonableness; it must also be supported by the right reasons, often by reasonable judgments of the legitimacy and relative urgency of the needs and interests of affected persons. Alexy's rules do not entail that only (potentially) justifiable reasons will be advanced: discourse ethics must rely upon a notion of relevance necessitated by $U$ in order to specify considerations it would be appropriate to advance in attempts to influence the outcome of the deliberative process.

A complete theoretical analysis of the idea of the reasonable may not be possible. Judgments about what is reasonable are not fixed and universal, and it makes sense to think that an analysis of the notion of the reasonable requires contextualization before it will yield any results. On the importance of contextualization, we should agree with Habermas. What it is reasonable to reject will often depend on purposes and values whose content and importance are socially relative. Moreover, moral discourse is required in order to work out the content of the reasonable within given contexts and with regard to particular cases. Whether our reasons are morally acceptable is something we determine in concert with others in view of the values and practices we share with them. There is no authority beyond agreement among participants in (ideal) moral discourse for determining what counts as a reasonable proposal. But we should reject the idea that reasonable procedural constraints on discourse can be understood to be purely formal. The role for substantive moral judgments about how discourse should proceed cannot be exhaustively represented by formal constraints on the discourse situation. In moral discourse we rely upon substantive ethical judgments to guide our principled assessment of what could count as a legitimate outcome. Such judgments are crucial if we are to be able reasonably to convince others that they have good reasons to endorse our proposals. The implementa-

---


67 Also relevant may be considerations of responsibility and fairness, see Scanlon, *What We Owe To Each Other*, chap. 5.

Habermas suggests that in stepping back to describe what he takes to be purely formal constraints on discourse (including $U$), the moral philosopher can retain an observer status. But a description of discourse ethical constraints is meaningful only in the context of discussing what is involved in applying them in practice. This requires the philosopher to engage in some substantive moral reasoning, however indeterminate and fallible it might be—that is, to engage as a participant in moral discourse. Further contextualization may well be required in order to resolve actual disputes but, nevertheless, the line between participant and observer has already been blurred. Moreover, since the requirements of reason cannot be codified by a list of formal constraints on a procedure of reasoning, the outcome of discourse can always be checked and possibly revised upon reflection. The outcome of actual moral reasoning is no more immune to examination and revision by reasoned reflection than is our best judgment of the outcome of ideal discourse.

4. Concluding Remarks on Justification

I have expressed doubt about whether discourse, broadly speaking, presupposes the norms articulated by Alexy’s rules. But I shall not pursue this matter further, for I do not agree that it is crucial to the idea of a discourse ethics that the justification of its principles rely on such a broad understanding of communicative action. It would be sufficient for discourse ethics to give a plausible and detailed account of how it is that

---

69Furthermore, the content of these judgments is not itself generated simply through our recognition of the validity of Principle $U$. For related discussion, see Barbara Herman, “Moral Deliberation and the Derivation of Duties,” in The Practice of Moral Judgment (Cambridge, Mass.: Harvard University Press, 1993), pp. 145-58. Herman argues that maxims of moral agents have implicit moral content; thus, moral judgment does not function only after maxim formulation (i.e., with the application of the Categorical Imperative).

70Habermas obviously thinks our attitudes as participants in moral relationships (i.e., our “performative attitudes”) are important to morality. (See his discussion of Strawson in Moral Consciousness and Communicative Action, pp. 45-50.) This is the very subject of his “formal pragmatics.” Nevertheless, in analyzing our performative attitudes Habermas suggests he is not at the same time engaging with them. See his distinction between the philosopher’s “reconstructive” focus on procedure and the engaged stance of a participant in discourse, in “Reconciliation through the Public Use of Reason,” p. 131.

71I do not wish to deny, however, that a broad conception of communicative action is crucial to the role of Habermas’s discourse ethics within his broader project of achieving a “rehabilitation of the concept of reason” to fit modern social practices and the potential they possess for an increasingly rationalized lifeworld. See his “The Normative Content of Modernity,” in The Philosophical Discourse of Modernity, trans. Frederick Lawrence (Cambridge, Mass.: The MIT Press, 1987), pp. 340-41.
communicative engagement with the normative regulation of our actions is central to social life, and how reflection on the nature and value of this activity leads us to affirm Principle $U$. Discourse ethics is not, as Habermas fears, in danger of relying upon a *petitio principii*. The idea that in making normative claims we are *doing* something to constitute our common life as *social*, is both plausible and illuminating. When we make moral claims on one another, we are establishing and re-establishing a society ordered by norms. Moral philosophy has been much preoccupied with the question of what our moral claims are *about* and has perhaps neglected the question of what our utterance of them *achieves*. When we advance and evaluate moral justifications, Habermas is right to stress that we are engaged in an important activity, one intrinsic to what it is to be in society with others.

Habermas advances arguments for the basic idea of discourse ethics from social theory, psychology, and semantics. He draws on social theory in order to promote the view that the nature of group membership supports the idea of communicative ethics. Communicative action, he argues, is intrinsic to the processes of the reproduction of cultural traditions, social integration, and socialization.

In coming to an understanding with one another about their situation, participants in interaction stand in a cultural tradition that they at once use and renew; in coordinating their actions by way of intersubjectively recognizing criticizable validity claims, they are at once relying on membership in social groups and strengthening the integration of those same groups; through participating in interactions with competently acting reference persons, the growing child internalizes the value orientations of his social group and acquires generalized capacities for action.

Under the functional aspect of *mutual understanding*, communicative action serves to transmit and renew cultural knowledge; under the aspect of *coordinating action*, it serves social integration and the establishment of solidarity; finally, under the aspect of *socialization*, communicative action serves the formation of personal identities. The symbolic structures of the lifeworld are reproduced by way of the continuation of valid knowledge, stabilization of group solidarity, and socialization of responsible actors. The process of reproduction connects up new situations with the existing conditions of the lifeworld; it does this in the *semantic* dimension of meanings or contents (of the cultural tradition), as well as in the dimensions of *social space* (of socially integrated groups), and *historical time* (of successive generations). Corresponding to these processes of *cultural reproduction, social integration, and socialization* are the structural components of the lifeworld: culture, society, person.

---

72The focus within ethical theory on arguments for and against realism, for instance, illustrates this point.

73Cf. Strawson, "Freedom and Resentment."


Discourse ethics claims we are committed to Principle D insofar as we are engaged within these dimensions of social life. This is interesting and compelling. Persons (at least some persons) must share some understanding of the norms of a cultural tradition and must accept them as valid if they are to sustain them. Social groups are strengthened and integrated when people's actions are coordinated by what they mutually recognize as valid norms. Socialization involves the internalization of roles and norms one takes to be legitimate. All these processes rely upon some grasp of the idea that valid moral norms should be generally acceptable, an idea discourse ethics analyzes.

Habermas also draws upon developmental psychology in order to argue that the requirements of discourse ethics—in particular, those that draw upon our ability to take up an impartial point of view and to act on principles—fit with our moral experience. And finally, his argument leans heavily upon claims about the meaning of normative concepts. Habermas writes: "In everyday life... no one would enter into moral argumentation if he did not start from the strong presupposition that a grounded consensus could in principle be achieved among those involved. In my view, this follows with conceptual necessity from the meaning of normative validity claims." His "transcendental pragmatic" arguments sometimes draw directly on claims about the meaning of normative concepts, especially to spell out when and how a pragmatic contradiction can be achieved.

All these considerations work within Habermas's theory of discourse ethics to show that ethical principles presupposed by what we do when we make normative claims find support in the nature of our social interaction and psychological development, and in the meaning of our concepts. These multiple sources of support play a crucial role in the justification of Principle D and, I claim, in the justification of U as well.

Habermas calls his account of discourse ethics "reconstructive." What he means by this is that while principles of discourse ethics are not necessarily consciously formulated by those (all of us) who in fact rely upon them, we can reconstruct their content in theory. Knowledge that the principles apply can be reconstructed from claims about the pretheoretic...
cal activities in which we actually engage; “knowing that” is reconstructed from a study of “knowing how” (i.e., from a study of successful regulative speech acts). As we have just seen, the reconstruction draws upon many sources, and involves attempts to balance the relevant considerations. All of this is not at all dissimilar to a philosophical clarification of the content of morality by way of an appeal to our judgments in “reflective equilibrium.” On a reflective equilibrium account, the burdens of justification or evidential support are carried (or may be carried) by our moral reflections at all levels of generality. Justification is generated by the value we ascribe upon reflection to a range of pretheoretical moral activities. It is of course important that Habermas’s reconstruction aims to test not just how we can best describe the content of our judgments, but how we should understand their role vis-à-vis our interactions with one another. What function do they serve? What does our concern with them achieve? Understanding this, I have argued, illuminates the value we place on the activities and relations that comprise the normative structure of social life. All of the many sorts of considerations (sociological, psychological, semantic) that help to elaborate these activities and what we value about them can fit within an ethical theory that claims to be justified in the end by appeal to nothing more foundational than our judgments in reflective equilibrium. They could, for instance, lend support both to the notion of moral motivation central to Scanlon’s account, and to the normative content of his idea of reasonable rejection. The “reconstructive sciences” offer more than what Habermas sometimes suggests is merely a check on a fallible “transcendental-pragmatic” deduction of $U$, they provide additional, independent sources of justification. In this way, the justification of $U$ can draw directly on the sources for formulating $D$.

$U$ represents a philosophically clearer and sharper statement of the content of morality indicated in Principle $D$: it zeroes in on the importance of fair and equal consideration of the interests of all persons to the justification of moral principles for the general regulation of behavior. Such fair and equal consideration is a generally acceptable way to reach agreement among persons who do not share a comprehensive ethical or religious doctrine, but do express genuine respect and concern for one

---

82 See Rawls, Political Liberalism, pp. 8, 28, 45.
84 See Habermas, Moral Consciousness and Communicative Action, p. 32.
another. As I have stressed, what the reflective process of philosophical justification adds in presenting a reconstructive approach as having the authority to justify $U$ is the claim that the activities that the implicit implementation of $U$ helps to organize are activities whose value we have no reason to doubt. We affirm them as central to our shared social life. Habermas should accept that there is no threat of vicious circularity in a reflective, reconstructive account of the nature of morality, even when moral philosophy does not present morality as grounded in anything else—including communication or rationality generally. In fact it is odd that Habermas insists such grounding of discourse ethics is necessary if claims about its principles are to be justified, for he rejects “first philosophy” and claims his philosophy is “post-metaphysical.” With regard to the theory of argumentation, for example, he writes:

I presuppose that a theory of argumentation must take the form of an “informal logic,” because it is impossible to force agreement on theoretical and moral-practical issues either by means of deduction or on the basis of empirical evidence. To the degree to which arguments are deductively valid, i.e., compelling in terms of logical inference, they reveal nothing substantively new. To the degree to which arguments do have substantive content, they are based on experiences and needs/wants that are open to various interpretations in the light of changing theories using changing systems of description. Such experiences and needs/wants thus fail to offer an ultimate basis for argumentation.\(^8^5\)

Habermas appears to be unclear about the sort of justification, if any, he is offering for the principles of discourse ethics. He proposes a deductive foundation for Principle $U$ in a theory of argumentation, but he also seems to deny the possibility of such a foundation. He says of his formal pragmatics, “This deductive concept of justification is obviously too narrow for the exposition of the pragmatic relations between argumentative speech acts.”\(^8^6\) In fact, in that passage he claims to have given up the aim of justification altogether. “The fact that there are no alternatives to these rules of argumentation is what is being proved; the rules themselves are not being justified.”\(^8^7\) In view of these remarks, Habermas’s insistence that Principle $U$ must be derived from more general (and non-ethical) claims about the structure of argumentation generally if it is to avoid the threat of a petitio principii is puzzling and points to the tension I have indicated in his view.

But, as I have been arguing, discourse ethics is not in need of any foundational support. Justification need amount to nothing more than the broad support the presuppositions of moral discourse receive in the nature of our social life as a whole. Reflection on the nature and importance

\(^8^5\)Ibid., p. 63.
\(^8^6\)Ibid., p. 79.
\(^8^7\)Ibid., p. 95.
of moral reasoning can bring together a wide variety of considerations, and justification is provided by their mutual support and overall plausibility. Of course, Habermas is interested in articulating a "universal core" of morality, and would shy away from a reflective equilibrium approach as unable to establish that the core claims of morality hold universally. It is true that attaining reflective equilibrium does not ensure that we will reach results that hold universally. The nature and presuppositions of communicative action associated with practical discourse might vary from one social context to another, just as I have argued they would likely differ from the presuppositions of theoretical discourse. But the proposed universality of core claims could be supported by evidence that the features of social life and human psychology that make the conception of morality expressed by Principle U possible and valued are universally present. A priori arguments for the idea that principles of discourse ethics hold universally are anyway unlikely to be convincing.

An analysis of the presuppositions of moral reasoning can be thought to succeed in justifying discourse ethics. It does so insofar as it is convincing as an analysis of an activity that is at the core of a social life in which what we recognize and affirm as moral relations with others are possible. Habermas offers a powerful analysis of the special status of practical discourse, for the aspects of social life he illuminates are aspects we value highly. Justification emerges from the mutual force of the many sorts of considerations that confirm this.

Erin Kelly
Department of Philosophy
Tufts University
ekelly@tufts.edu

---

88 Communicative and strategic action may be harder to separate on some conceptions of morality. On others, justification that aims for agreement may be addressed to a select group only. See Rorty, Contingency, Irony, Solidarity, and also his Objectivity, Relativism, and Truth: Philosophical Papers, Volume I (Cambridge: Cambridge University Press, 1991).

89 Nevertheless, it is possible that some societies could find that something other than persons’ reciprocal affirmation of each other’s interests—such as deepening shared devotion to a particular religion or culture—will bring about the possibility of agreement.

90 I am grateful for the guidance I received in writing this paper from Joshua Cohen, John Rawls, and T.M. Scanlon. I thank the members of the Moral and Political Philosophy Workshop at Harvard and my colleague Hugo Bedau for their constructive criticisms. This paper was also improved by discussions with Hilary Putnam and a close reading by David Peritz. Finally, I would like to thank the anonymous referees for Social Theory and Practice for their very helpful suggestions.